AMERICAN JOURNAL of PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 19, NUMBER 5 MAY, 1951

The Employment of Physicists by the Department of the Army*

MARSH W. WHITE†
The Pennsylvania State College, State College, Pennsylvania
(Received November 22, 1950)

Physicists are being employed in installations of the U. S. Army in increasing numbers. Although research and development activities are their major function, many are engaged in administration, technical writing, liaison duties, and staff work. In many respects the work of physicists in Army laboratories is similar to that in typical university centers. Salary scales are generally higher than those for comparable academic positions but below those often paid in industry. Many positions filled by physicists are classified by titles not clearly identified with physics. Physics training as a background field is a requisite by many Army scientists who are not professional physicists. The Army facilities for experimental work in physics are excellent. Most Army research and development centers are administered by the technical services. Army policy provides that scientific employees be given opportunity for considerable professional development through publication of technical papers, attendance at scientific meetings, and further study.

THE nature of wars as presently fought and as visualized in the future requires that applied science be utilized in ever-expanding degree. Because of the contributions of physics and physicists in World War II, the current planning of the Department of the Army contemplates strong emphasis upon physics in its research and development activities. Although the Army has traditionally employed civilian physicists for creative work in research and development and as administrators and directors of activities in various other fields, present conditions call for the employment of considerably larger numbers of physicists of varying degrees of training and experience. Furthermore, increasing numbers of Army officers are being educated in civilian graduate schools and taking advanced degrees in physics.

Army research and development budgets are large, in comparison with the corresponding prewar appropriations. There seems to be good reason to presume that this situation will continue during the predictable future. The bulk of the actual laboratory work done in Army installations is assigned to civilians, with military personnel in positions having administrative and executive responsibilities. Many of these administrators have graduate degrees in the physical sciences. Because of the shortages of competent scientific personnel, some Army installations have not been able to recruit all of the physicists desired during certain portions of the postwar interval. During current and future emergency conditions it is expected that these

* An article dealing with this subject is also to appear in a forthcoming Civil Service Commission pamphlet entitled *The Physicist in the Federal Civil Service.*

† Consultant to the Department of the Army, Research and Development Division, Assistant Chief of Staff, G-4 (Logistics).

shortages will be materially increased. Almost all branches of physics are covered in Army requirements.

While research is one of the best-known functions of physicists in Army laboratories, many perform other functions. These duties include consulting, testing, designing, administration, writing, and liaison with industry, universities, and other research organizations. Often the major projects on which a physicist is working require participation in a number of these types of activities. So many of these functions overlap that it is difficult to establish clear lines of demarcation between them. It is therefore apparent that many research physicists who desire employment in Army installations must be qualified with certain personal characteristics other than ability as theorists or as laboratory experimentalists.

The proportion between *research* as distinguished from *development* in Army laboratories varies greatly from one project to another. It is often difficult to specify the degree to which a project should be classified as research or as development. Some laboratories resemble typical university research centers. Others are ostensibly devoted only to the development of weapons for warfare. Even in the latter case, however, many projects require work which might properly be classified as research.

Research may be divided into two major categories: basic (or pure) and applied. A third classification, background research, is often used. All of these types of research are common in Army installations. It is easy to see why most Army research projects involve work in all three of these types of research. Salaries, promotion, and opportunities for self-improvement are not especially different for physicists engaged in work in any of the three types of Army research.

Many positions filled by physicists in Army installations are not classified by a title which clearly identifies this science. A number of job titles are designated by such terms as "physical scientist, technologist, radiologist, radar technician, electronics engineer," and a variety of other occupational specialties. Those jobs of an administrative, supervisory, editorial, or educational nature are usually classified by titles not easily identified with physics. For these reasons it is difficult to state with any exactness the numbers of civilian physicists employed by the Army. It is certain that the number is measured in the hundreds. This is exclusive of the several hundred junior physicists engaged each summer as temporary student-aid trainees.

The minimum Civil Service grade of regularly employed Army physicists is GS-5, with a beginning annual salary of $3,100. The grades extend upward, with a considerable fraction employed in the highest grades. A number of Army physicists hold the high grades of GS-16, 17, 18, and P-9, the so-called "super" grades, with salaries of $10,000 to $15,000. Recent studies show that the median salaries of physicists employed by the government are higher than those paid by educational institutions; they compare favorably with the salary scale in industry. Although the Army cannot compete with industry on a salary basis for the services of some of the more eminent scientists, it has procured their part-time services in many instances as consultants in the solution of spot problems.

Most Army laboratories employ many scientists, engineers, and technicians who are not professionally trained primarily as physicists but who are expected to have a considerable background in this science. This group includes all of the usual fields of engineering, chemistry, metallurgy, the earth sciences, and the borderline subjects such as biophysics, physical chemistry, and some of the medical sciences.

Because of the considerable numbers of physicists employed, the expansion of certain activities and the variety of the work to be done, it is currently true that unusual promotional opportunities exist for qualified physics personnel in Army installations. This is particularly true in the lower and upper grades. Like many government agencies the Army follows a system of making the majority of its promotions from within the ranks of its civilian employees. When vacancies occur or when new positions are created, the qualifications of present employees are carefully scanned in attempts to locate the desired personnel from the ranks of the existing employees.

Many Army laboratories have outstanding facilities for scientific research and development. The government has appropriated adequate

funds in recent years to enable the Department of the Army to equip its technical installations with almost every conceivable device and apparatus necessary to accomplish their missions. While much of such equipment is typical of that found in University laboratories, a great deal is specially designed and often unique. Typical of the latter class would be such devices as the electronic computers, high-speed cameras with exposures in microseconds, systems for transmitting data rapidly over long distances, unusual biophysical research plants, ultrahigh frequency radar sets, and machines for studying materials and equipment under a wide variety of extremes of temperatures, pressures, humidities, etc.

The employment of physicists by the Army is widely distributed over the country. Only a few work in Washington, as there is no concentration of Army laboratories in the District of Columbia. Most of the research and development work which utilizes physicists is performed by the Technical Services of the Army. These, together with their major research installations using physicists, are as follows:

Chemical Corps

Edgewood Arsenal, Maryland (Army Chemical Center)
Camp Detrick, Frederick, Maryland

Corps of Engineers

Engineer Research & Development Laboratories, Fort Belvoir, Virginia

Medical Corps

Walter Reed Hospital, Washington, D. C.
Medical Nutrition Laboratory, Chicago, Illinois

Ordnance Corps

Aberdeen Proving Ground, Aberdeen Maryland
Ordnance Rocket Division, Fort Bliss, Texas
White Sands Proving Ground, Las Cruces, New Mexico
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California
Frankford Arsenal, Philadelphia, Pennsylvania
Watervliet Arsenal, Watervliet, New York
Watertown Arsenal, Watertown, Massachusetts
Picatinny Arsenal, Dover, New Jersey
Redstone Arsenal, Huntsville, Alabama
Rock Island Arsenal, Rock Island, Illinois

Quartermaster Corps

Quartermaster Food & Container Institute, Chicago, Illinois

FIG. 1. A no-echo silent chamber designed by members of the Coles Signal Corps Laboratory staff. This anechoic chamber is being used to study microphones, head sets, loudspeakers, and other acoustical devices. (U. S. Army Photograph.)

Climatic Research Laboratory, Lawrence, Massachusetts
Philadelphia Quartermaster Depot, Philadelphia, Pennsylvania

Signal Corps

Squier Signal Laboratory, Fort Monmouth, New Jersey
Coles Signal Laboratory, Red Bank, New Jersey
Evans Signal Laboratory, Belmar, New Jersey

In addition to the above there are a number of minor laboratories and also other Army agencies which employ scientific and technical personnel. For example, a number of physicists act as technical specialists and consultants in various staff groups concerned with Army research and development. Some physicists serve abroad with the occupation forces of the Army.

As examples of Army installations employing physicists some of the work of interest done in the Signal Corps and the Ordnance Corps will be briefly outlined. The Evans Signal Laboratory is the center for work on radar, meteorology, thermionics, vacuum tubes, and devices for the protection and detection of sound, heat, and light. At the Coles Signal Laboratory research and development proceeds in radio, telephone, and communication systems, in television, acoustics (see Fig. 1), and electrical power devices. The Squier Signal Laboratory is the focal place in the Army for research on frequency control and quartz oscillator apparatus; here broad

FIG. 2. Flexible-throat wind tunnel in operation at the Ballistics Research Laboratory, Aberdeen Proving Ground. A guided missile is being tested at Mach number 4. (U. S. Army Photograph.)

studies are conducted on photography by motion and still cameras. Here also work is done on batteries, component parts, circuit elements, and internal combustion power machinery. The three Signal Corps laboratories range over an area of 300 acres, providing test sites for any activity from mobile television to radiosonde flight.

As an example of the exceptional facilities of a Signal Corps installation those at the Thermionics Branch Electronic Laboratory are listed in brief detail: complete facilities for CW or pulsed life test, simultaneously testing 1000 tubes under any operating conditions on frequencies up to 50,000 megacycles; high power substations for peak pulse operation up to 50 kw and to 100 kw for continuous power ratings; spectrum analyzers for all frequencies; complete equipment for the study of television, cathode-ray and phototube behavior, including focusing, resolution, and photosurface characteristics; fully equipped glass-blowing laboratory with hydrogen furnaces, induction bombarders, pumping stations, and bake-out furnaces; extensive solid-state facilities for the investigation of transistors and similar devices.

In the Ordnance Corps, the physicist enjoys a stimulating variety of research opportunities. His career has the benefit of a stable and enlightened research policy guided by an Advisory Committee of the Nation's leading scientists. He has at his disposal an abundance of excellent research facilities backed by the huge technological resources of the great Arsenals and Proving Grounds.

Ordnance research includes substantially the whole range of physics, theoretical and experimental, fundamental and applied. As a developer of Army weapons, the Ordnance Corps conducts research in electronics, optics, photography, spectroscopy, nucleonics, the solid state, surface phenomena, supersonic and diabatic flow, and numerous other fields. There are also activities in the borderline fields, such as chemical physics, astronomy, and meteorology.

One example of an Ordnance research institution is the Ballistics Research Laboratories at Aberdeen Proving Ground, Maryland. These laboratories have the air of an active university physics department. An outstanding research staff works in modern, well-equipped laboratories (shown in Figs. 2 and 3) and offices on a typically wide range of modern physics problems. The Computing Laboratories have the famous electronic and other computing "brains" in a special wing built to house them. There is an excellent, up-to-date scientific library including complete files of classified Government documents; every phase of modern scientific literature is available to the Ordnance physicist.

At the weekly Ordnance Colloquium, staff members present papers on subjects as varied as astrophysics and the mathematical theory of games. The young and growing Ballistic Institute offers specialized courses at graduate level and represents one phase of the Ordnance policy to promote the training of scientific personnel. Typical of Ordnance laboratories, there is nearby a large metropolitan center, in this case Baltimore. A number of Ordnance personnel, while

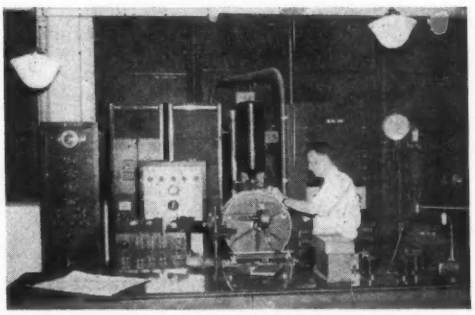


FIG. 3. A physicist at the Ballistics Research Laboratory, Aberdeen Proving Ground, operates the high speed drum camera of a specially designed cathode-ray oscillograph. (U. S. Army Photograph.)

working at the laboratories, have continued their graduate studies at Johns Hopkins University, the University of Maryland, or the University of Delaware.

As at all Ordnance installations, there are frequent occasions for scientists to travel to other installations for consultation or work. For example, the Aberdeen laboratories have an Annex at the White Sands Proving Ground, where the Ordnance Guided Missile firing program is conducted (see Fig. 4). There is a continual flow of Ordnance personnel to and from White Sands for the purpose of conducting experiments and tests in connection with the missile firing program.

The Ordnance laboratories enjoy frequent visits and lectures by top American and foreign scientists. The general policy with respect to the research program of the Ordnance Corps laboratories is guided by a distinguished Scientific Advisory Committee. This committee includes the following: Francis H. Clauser, Theodore von Kármán, J. von Neumann, I. I. Rabi, H. N. Russell, and H. C. Urey. At several meetings each year, Ordnance Corps scientists meet with the committee and enjoy individual contact with its members in discussion of research problems.

The physicist in some Army laboratories may follow the development of his own embryo idea to completion of the final working product. He begins by experiments, supplemented by theoretical computations. Then laboratory models are prepared and laboratory tests are conducted. After the preliminary designs are perfected, working models are made and field tests are conducted. The experimenter frequently has the satisfaction of seeing his work result in production and maintenance data necessary for the widespread use of his device in the theatres of operation.

It is the policy of the Army Technical Services to keep their professional employees acquainted with the activities which should make them familiar with the recent developments in their fields. The Army also gives to its scientists opportunity for professional development through contact and information exchange with the leading authorities in their fields. Publication of the results of unclassified research in technical journals is facilitated. Encouragement and opportunity to affiliate with and to attend the meetings of scientific societies is given. Opportunities for advanced graduate study are provided at many Army laboratories. Graduate study programs are usually operated in connection with nearby universities. Classes are frequently held within the laboratories at times convenient to the students. When the course of study is determined to be directly related to work being performed, the Department of the Army frequently provides time off for attending class during working hours and may contribute toward the cost of tuition.

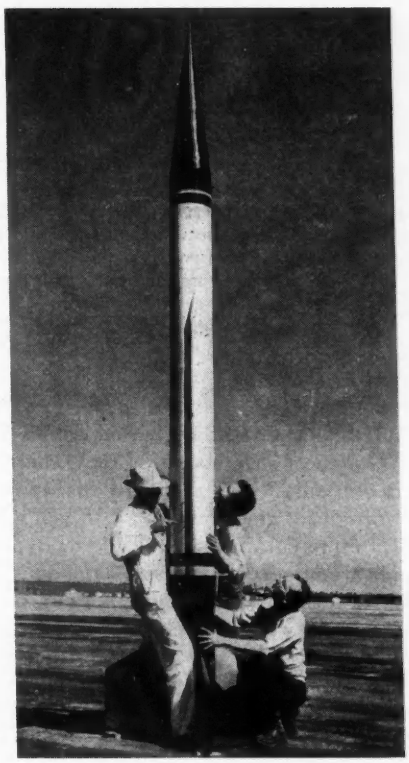


FIG. 4. A "WAC Corporal" rocket being prepared for test at the White Sands Proving Ground. (U. S. Army Photograph.)

Practically all of the positions for civilian physicists in the Army Establishment are under the Civil Service and Classification Acts.

Potential against Resistance—a Graphical Review

W. W. Sleator
University of Michigan, Ann Arbor, Michigan
(Received June 21, 1950)

In this paper the graph of potential against resistance, often used to picture the discharge of an electric cell, is shown to represent the transformations of energy involved in the process. With this property in view, the $V-R$ diagram becomes applicable to every energy transformation that can occur in an electric circuit. Such a diagram is used accordingly to represent the charging of a storage cell as well as its discharge, the charging of a condenser, the operation of the dc generator and motor, the performance of the copper-iron thermocouple, and the increase and decrease of current in an inductive coil without iron. It is hoped that this unified treatment of important electric processes may reveal their essential similarities and make them better understood and more easily remembered.

J. WILLARD GIBBS begins his famous paper, "Graphical Methods in the Thermodynamics of Fluids"[1] with a sentence that may very well introduce this review. "Although geometrical representations in the thermodynamics of fluids are in general use, and have done good service in disseminating clear notions in this science, yet they have by no means received the extension in respect to variety and generality of which they are capable." The same situation exists in the conventional treatment of electricity in general physics, where a graph of potential against resistance is used in many books to represent the discharge of a cell or battery through internal and external resistance, but no other use is made of these coordinates. The fact is that a graph in a potential-resistance coordinate system illustrates the charging of a storage battery as well as its discharge or use, it exactly describes the process of charging a condenser, it clearly presents the most essential features in the operation of the direct current motor and generator, it gives a good account of the performance of the thermocouple, and it represents so well the effect of self-inductance as to clarify notably Faraday's discovery and Lenz's law. In these cases the figures make it easy to follow the fundamental process, namely, the transformation of energy. They suggest and exemplify a principle, applicable throughout an elementary treatment of electricity, to which there are no exceptions. In the absence of such a principle any general account of electricity is a collection of disconnected facts.

I. USE OF A STORAGE CELL

We begin with the familiar figure that represents the discharge or use of a storage cell of internal resistance r and emf E. In order to employ the conventional current we assume that positive electricity only is in motion through the circuit, and arrows on the graphs indicate the direction in which electricity, considered as positive, is moving. If it is preferred to follow the motion of negative electricity, the arrows indicating direction on these figures may be reversed. In all these figures we use V for potential and *Res.* for resistance in general, with r and R for resistances internal and external to the cell or other device considered.

The connections are shown by Fig. 1(A), where point O next to the negative electrode has the lowest potential in the circuit. The graph of potential against resistance is given in Fig. 1(B). Here $OA=r$ and $AO'=R$, so that in the circuit O' is the same point as O. The ordinate at any point in the lines JK and PO' indicates the potential (with respect to O) of the corresponding point or intersecting surface in the circuit. Accordingly, the graph $OJKPO'$ is the path, in the resistance-potential plane, of any element of positive charge in its migration around the circuit. The directions indicated make it obvious that electricity, considered as positive, goes from higher to lower potential where resistance only is encountered, and there electrical energy disappears and heat is developed. It goes from lower

[1] *The Scientific Papers of J. Willard Gibbs* (Longmans, Green, and Company, New York, 1906), Vol. 1, *Thermodynamics*, p. 1.

potential to higher where the electrodes are in contact with the solution, and there electrical energy appears at the expense of chemical energy. It is evident that lines JK and PO' must be parallel, and their slope is the current, so that $\tan\theta = I$.

In Fig. 1 we have supposed that the single potential differences OJ and KP are known, and that each represents a rise in potential as we go counterclockwise around the circuit. However, the quantities we are able to measure directly are E, E', r, and R. Hence, we may better represent the circuit as we know it by Fig. 2. Here any one of the dotted lines may show the falling potential within the cell, and each of these inclined lines gives the same E and the same E'. If we are representing a Daniell cell and are convinced that the copper electrode is negative

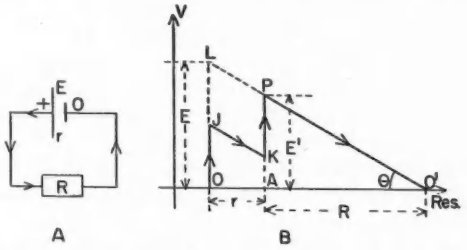


FIG. 1. Discharge of a cell: (A), the simple circuit; (B), graph of potential against resistance for the complete circuit; R and r are the external and internal resistances, respectively.

with respect to the solution, we will use the upper incline, J_2K_2. But Fig. 1 indicates well enough the essentials of the discharging process.

The graph of Fig. 1 is readily adapted to the short-circuited cell, and to the open circuit also, as shown in Figs. 3(A) and 3(B). For in Fig. 3(A) the external resistance is zero, hence $E' = 0$, and in Fig. 3(B) the current I is zero and $E' = E$. There are no arrows on the graph of Fig. 3(B) because there is no motion of electricity in the conductors.

We now state the general principle mentioned in the first paragraph.

A. In that part of a circuit where electricity, considered as positive, is moving upward in potential, so that the motion of any element of positive charge carries it from a position of lower to one of higher potential, energy appears in electrical form, or electrical energy appears, and energy in some other form or forms disappears in equal amount.

B. In any part of a circuit where electricity, considered as positive, is moving downward in potential, so that the motion of any element of positive charge carries it from a position of higher to one of lower potential, electrical energy disappears, and energy in some other form or forms appears in equal amount.

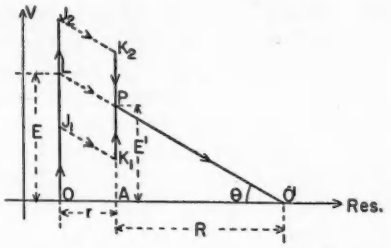


FIG. 2. Discharge of a cell: the potentials E, E' and the resistances R, r are presumed known. Additional information is needed to determine whether the discharge is truly represented by the path that includes J_1K_1 or by that which includes J_2K_2.

We can obviously state the principle in terms of negative electricity by interchanging the words upward and downward, and higher and lower. The other kinds of energy alluded to here make up all the kinds of energy there are, except electrical. We shall illustrate the appearance and disappearance of heat and chemical energy (already mentioned in dealing with the use of a cell), mechanical energy, and the energy of the magnetic field. We shall be dealing with the simpler situations in which the current is the

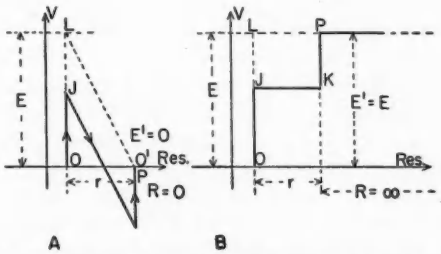


FIG. 3. Special cases of simple electrical circuits that include a cell and an external resistance: (A), short circuit, for which $E' = 0$; (B), open circuit, for which $E' = E$.

same, at any one time, in different parts of a single circuit. The current i, defined as dq/dt, is the rate at which electricity crosses any surface intersecting the conductor. We find in a recent book the statement that "When an electric current flows, charge is transported from a place of high potential to a place of lower potential." Principles A and B assure us that if this is the case in any part of a circuit, charge must be transported from a place of low to a place of higher potential in some other part, and guard us against the idea that electricity always "runs down hill." Students must not lose sight of the fact that in tracing a complete circuit, one must encounter rises in potential that are together equivalent to the total fall.

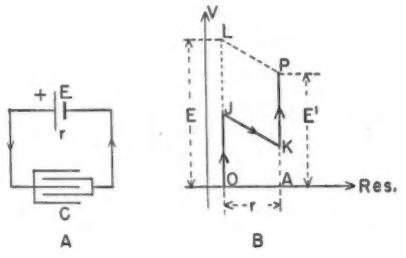


FIG. 4. Charging a condenser: (A), the simple circuit, omitting switches; (B), the corresponding resistance—potential diagram. The internal resistance r of the cell is presumed to be the total resistance of the circuit whose self-inductance is taken to be zero.

II. CHARGING A CONDENSER

We now consider a cell connected directly to a condenser originally uncharged. Figure 4(A) gives the circuit and Fig. 4(B) the resistance-potential diagram. We take the circuit to have the resistance r and self-inductance zero. Of course, there will be a little self-inductance in any circuit; but if r is not too small, the current will rise to its maximum value in a very small part of the time required for it to fall. Before the cell is connected, the terminal potential difference is its emf, as indicated in Fig. 3(B); but just after the connection is made, and before an appreciable amount of electricity is transferred, the situation is represented by Fig. 3(A). Then, as time passes, conditions change from those of Fig. 3(A) to those of Fig. 3(B), and when the condenser has about $\frac{3}{4}$ of its final charge, the existing potentials are shown in Fig. 4(B). Here E' is the potential difference between the terminals of the cell and also that between the terminals of the condenser. Under the circumstances of Fig. 4(B) consider the passage of the time interval 10^{-10} sec, for example, during which E' does not appreciably change. In this interval, a small charge, Δq, crosses each section of the metallic and liquid parts of the circuit, so that Δq leaves the negative side of the condenser, and an equal charge Δq (but not the same electricity) arrives at the positive side, potential E' higher. Charges and potentials throughout the circuit being unchanged, the effect is the same as if Δq had been transported from one side of the condenser, through the cell, to the other side, having been carried up the net amount E' in potential. If the energy required for this transfer is ΔW, then $\Delta W = E'\Delta q$, and this amount of energy has been contributed by the cell. If now we plot the terminal potential difference E' as ordinate against the charge q in the condenser as abscissa, we have the straight line of Fig. 5, for the ratio q/E' or q/V is the constant capacity C of the condenser. Then the area of the narrow strip above Δq is approximately $E'\Delta q$, and this is ΔW. The limit of the sum of all such areas is the area of the triangle OPQ, and this is the limit of the sum of such terms as ΔW, and therefore is the energy of the charged condenser when $E' = E$. Therefore,

$$W = \tfrac{1}{2}EQ = \tfrac{1}{2}CE^2 = \tfrac{1}{2}Q^2/C.$$

In these expressions, if we use E in volts, Q in coulombs, and C in farads, W is given in joules. If we write $dW = (q/C)dq$ we can integrate, obtaining $W = \frac{1}{2}Q^2/C$ as a check on the graphical method.

Before going on to motors and generators we note that, whether the resistance of the circuit is inside the cell or outside, or partly one and partly the other, just one-half the energy lost by the cell is transferred to the charged condenser, and half is converted into heat by the resistances. We have a check on this result, which is beautiful in itself even if no check is required. Let the current at any time during the charging process be i, so that $i = dq/dt$. Then $q = C(E - ir)$, if r is the total resistance of the circuit. This gives a differential equation for q in which the variables are separable at once and which leads

to the equation $q=EC(1-e^{-t/rC})$. Accordingly, $i=(E/r)e^{-t/rC}$. Then for the total heat H appearing in r, no matter where r may be, $H=\int i^2 r dt=\frac{1}{2}CE^2$. But this is just the same as the energy of the completely charged condenser, and just half the total loss by the cell, confirming the former result. The reader who has not already carried out the details of this check will be well repaid for doing so.

III. CHARGING THE STORAGE CELL

In order to charge a storage cell, the necessary generator (of whatever kind) must lift electricity, considered as positive, to a potential high enough so that it will move down hill, or down in potential, through the cell. Figure 6(A) shows the circuit, where the resistance r_1 is in the generator, r_2 and r_4 serve merely to separate the generator a little from the cell, and r_3 is in the cell. Figure 6(B) shows the graph of potential against resistance as before. Comparing Fig. 6 with Fig. 1, we see directly that

In the use or discharge of a cell by a current i_1,

$$TPD=E'=E-i_1r,$$

so that the terminal potential difference (TPD) is necessarily less than the emf, and that in the charging of a cell by a current i_2

$$TPD=E'=E+i_2r,$$

so that the terminal potential difference is necessarily greater than the emf. We find that the actual emf E of the cell during the charging process is somewhat greater than that during use, and this difference reduces the efficiency of the charge-discharge cycle. But the essential difference between the two terminal potentials is due to internal resistance, and this is principally responsible for a possible low efficiency, which will obviously be less as the currents used are greater.

The term *counterelectromotive force* may properly be introduced in our account of the charging process. Electromotive force is defined as work per unit quantity of electricity. The emf of a cell in use is the potential energy lost by the cell per unit quantity of electricity crossing any section of the circuit, or branch of the circuit,

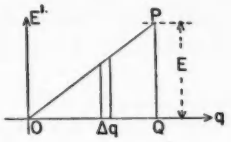


FIG. 5. Graph of the terminal potential difference E' of a condenser as a function of its charge q during the charging process.

containing the cell. Since neither work nor quantity of electricity has any intrinsic direction, we must set up an arbitrary definition of the direction of an emf. The direction of an emf is the direction of the conventional current produced. We agree that the part of a circuit where positive electricity is moving from lower to higher potential is the seat of an emf, its direction being that in which positive electricity is moving in that part.

Now in Fig. 6, electricity, considered as positive, is going down hill through the cell, that is from higher potential to lower. Hence, the current is opposed to the emf of the cell. By maintaining this countercurrent we do not reverse the cell, on the contrary, we are augmenting its natural polarity. Conventional language has a counter-emf in the cell under these conditions, but it might be better to speak of an emf, simply, and a countercurrent. Whatever the language we must see clearly that at a counterelectromotive force we do not lift electricity over an obstacle, we have only to allow it to fall, having lifted it enough by a generator in some other part of the circuit.

If the quantity of electricity Q crosses any section of the circuit of Fig. 6, the product EQ gives the amount of energy stored in chemical form in the cell, and $E'Q$ gives the total energy required. The difference is ir_3Q, which is the amount of heat appearing in the cell. It will be understood that by the use of a 220-volt dc

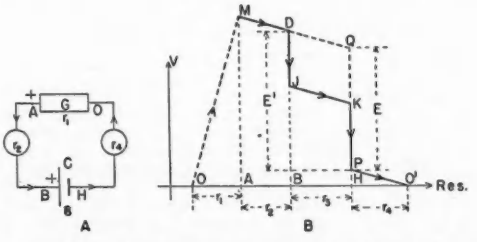


FIG. 6. Charging a storage cell: (A), the circuit including a generator G, a cell C to be charged, and resistances r_1, r_2, r_3, r_4; (B), the corresponding graph of potential against resistance. The line segments MD, JK, PO' are parallel.

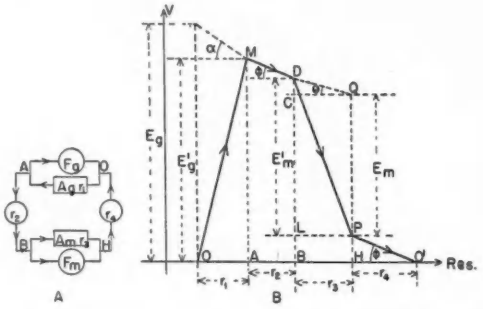


FIG. 7. Variation of potential against resistance for a circuit consisting of a shunt generator connected to a shunt motor: (A), the circuit; (B), graphical representation of the potentials at different points of the circuit. The line segments MD and PO' are parallel.

generator, for example, one might be able to send so large a current through a cell from the negative to the positive terminal as to reverse the natural polarity, because, while such a current was maintained, the product ir would be greater than the emf. Then positive electricity, moving in the direction of an emf, would be moving from higher potential to lower. But this need not disturb the definition agreed upon above, for here the motion of electricity is not caused by the emf. It is obvious that throughout this paper the term direction indicates only one way around a circuit, or the other. We have taken care to specify the meaning of the direction of an emf because most general physics books seem to assume that the student will know what this direction is without being told, which does not prove to be the case. The term counter-electromotive force will be exemplified in the following parts of this paper. In Fig. 6, we have represented the upward passage of positive electricity through the generator by a dotted line, since it was not necessary to specify any particular kind of generator.

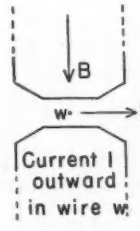


FIG. 8. Scheme of lecture experiment to make clear the distinction between motor and generator; the arrow B gives the direction of the magnetic flux; a current toward the reader is maintained by an outside source in wire W, which therefore experiences a force F. If the wire is allowed to move in the direction of F, the arrangement is a motor; if the wire is forcibly moved in opposition to F, it is a generator.

IV. DIRECT-CURRENT MACHINES

In Fig. 7(A) we have connected a shunt generator and a shunt motor, indicated by subscripts g and m, armature and field being lettered A and F. The potential-resistance diagram, Fig. 7(B), makes obvious the fundamental energy relations involved. The fact that $\alpha > \phi > \theta$ indicates the corresponding relation among the currents, for the current in the armature of the generator is greater than that in the line, that is in r_2 and r_4, and this current, in turn, is greater than that in the armature of the motor. If the quantity of electricity Q crosses a section of A_g, $E_g Q$ gives the mechanical energy that must be delivered to the generator on all electrical accounts, and if the quantity Q_1 crosses a section of the line, $E_m' Q_1$ represents the total electrical energy delivered to the motor. Also, $E_g > E_g'$, but $E_m < E_m'$. We see that the generator is closely analogous to the storage cell in use and the motor to the storage cell being charged. The essential difference is that with the machines, mechanical energy takes the place of the chemical energy in the cell.

A very simple lecture experiment makes clear the fundamental difference between motor and generator, and their essential similarity. We arrange an electromagnet so that there is a strong vertical field between the poles, as in Fig. 8. An extra long wire is used for one of the connectors, so that it can be brought between the poles, and the magnetizing current I is outward in the figure. The force F on the wire is directed toward the right. If the wire is allowed to move in the direction of F, the outfit is a motor. If the operator makes it move in opposition to F, that is, to the left, it constitutes a generator. That is just the difference. Now when the wire is allowed to move the motion to the right, and the motor action, are obvious. But when the wire is compelled to move to the left, the required push is obvious enough, but the current generated is imperceptible. We can, however, by use of a galvanometer and a few switches, separate the generated current from the magnetizing current so that it can be recognized.

It is stated in nearly all general physics textbooks that the emf generated in a motor (the counter-emf E in Fig. 7), multiplied by the arma-

ture current, gives the rate at which mechanical energy appears or is developed in the motor. This is correct, and of this energy not all, but commonly the larger part, is delivered to the mechanical load. The common explanation, however, if any, is about as follows. There is a fall in potential in the armature due to its resistance (DC in Fig. 7(B)) on which account some energy appears in the armature as heat, at the rate $i_a^2R_a$. But the rest, or major part, of the fall (CL in Fig. 7(B)) does not indicate the appearance of heat, and therefore, since there is nothing else it can give, it must represent, when multiplied by the armature current, the rate of appearance of mechanical energy. This argument by default is an unsatisfactory proof. We shall now give, in a single operation, an actual and direct proof that $(Ei_a)_g$ gives the rate at which mechanical energy must be supplied, on all electrical accounts, to the generator, and $(Ei_a)_m$ gives the rate at which electrical energy is converted into mechanical energy in the motor. This proof depends upon the fact, made obvious by the experiment of Fig. 8, that there is a single expression for the force on a wire carrying a current across a magnetic field, whether the wire is moved by that force, as in the motor, or by an opposite but equal force, as in the generator.

In either case we have an armature rotating in a magnetic field, mostly radial, and the familiar expression for generated emf applies as well to the motor as to the generator. It is

$$E=(\phi_1 Pnz)/(10^8 p).$$

Here ϕ_1 is the flux from one pole (maxwells or lines), P is the number of poles, n the number of revolutions per second, z the total number of wires parallel to the axis on or near the surface of the armature, and p the number of equivalent parallel paths through the armature, each containing z/p conductors in series. Quantities P and p are frequently the same, but P may be 4 or 6 while p is only 2.

Using i_1 as the current in one wire of length l, which is distant r from the axis, and B as the induction in the air gap, we have, for the force on one wire

$$f_1=Bi_1l/10 \text{ dynes},$$

the other units being gausses, amperes, and centimeters. In either machine then, as we have explained, the torque given by one wire is t_1 and

$$t_1=Bi_1lr/10 \text{ dyne cm}.$$

Then the number of wires actually in the magnetic field, assumed to be radial beneath the pole faces, is the fraction f of the whole number z, that is, fz. Also, the induction B is given by $P\phi_1/2\pi rlf$. For the combined area of all the pole faces is $2\pi rlf$, the area of the cylindrical surface of the armature being $2\pi rl$. Let the mechanical power be P'. Then $P'=T\omega=2\pi nT$. Here T is the total torque and ω is the angular velocity of the armature. Hence

$$\begin{aligned} P' &= t_1 fz 2\pi n \\ &= (Bi_1 lrfz 2\pi n)/10, \\ &= (P\phi_1 i_1 lrfz 2\pi n)/(2\pi rlf 10). \end{aligned}$$

But $i_1=i_a/p$; that is, the current in one wire is the armature current over the number of parallel paths, so that $P'=i_a(P\phi_1 nz)/(10^8 p)$ watts $=Ei_a$ watts, which was to be proved. This is the rate at which the motor develops mechanical work, or the rate at which mechanical work is delivered to the generator, but only on all electrical accounts. The motor actually supplies its load (saw, for example) at a rate somewhat less than this, and we have to supply the generator at a rate somewhat greater.

To both machines we can now apply Kirchhoff's laws. In the generator, $i_A=i_L+i_F$, where L represents the power lines and F the field, so that mechanical power supplied on all electrical accounts is

$$Ei_A=(E'+i_Ar_A)(i_L+i_F)=E'i_L+E'i_F+i_A^2r_A.$$

The three terms represent, of course, the rate of delivery of energy to the lines, the rate of appearance of heat in the field, and the rate of appearance of heat in the armature.

In the motor, $i_L=i_A+i_F$, and the electrical power supplied on all accounts to the motor is

$$E'i_L=(E+i_Ar_A)(i_A+i_F)=Ei_A+E'i_F+i_A^2r_A,$$

where the three terms represent, respectively, the rate of development of mechanical energy, the rate of appearance of heat in the field, and the rate of appearance of heat in the armature.

Here we have taken account of electrical losses only, but if in addition to these we know, for

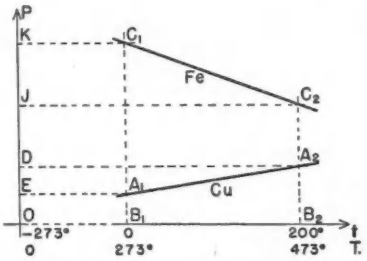


FIG. 9. Graph of thermoelectric power P for copper and for iron, each with respect to lead, as a function of temperature T.

example, that 0.85 of the mechanical energy developed in the motor is delivered to the load, we can calculate the over-all efficiency of the motor. If we know that the power delivered to the generator on all electrical accounts is, for example, 0.9 of the rate at which energy must be delivered to the machine on all accounts, we can calculate the over-all efficiency of the generator.

There is no use in trying to change an established custom that is in no way wrong, but it seems that the emf in a motor is no more "counter" than the emf of a generator, for they are similarly related to motion and magnetic induction. The current in the armature of the motor is counter to the current that the emf of the motor would produce. The term counter-emf seems to suggest an obstacle to be overcome, as if the positive electricity had to be pushed up hill when it got there. When principles A and B are applied, it is obvious that such is not the case.

V. THERMOCOUPLES. ELECTRICAL ENERGY AND HEAT

We have seen that a transformation of electrical energy into heat occurs when there is current in electrical resistance and that a resistance cannot be compelled to convert heat into electrical energy. The transformation conforms to general principle B (p. 263). However, the simple junction of two metals, and also a homogeneous conductor in which there is a temperature gradient in the line of flow, are devices by which the transformation is reversed when the current is reversed. If current through a junction, from metal A to metal B, gives rise to heat, beyond that due to resistance, the junction is the seat of a Peltier emf π directed from B to A.

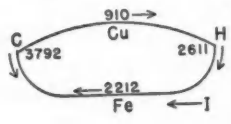


FIG. 10. The component thermal emfs (in microvolts) in a simple copper-iron circuit. The hot and cold junctions H and C are maintained in this case at 0° and 200°C, respectively.

We will consider a circuit consisting only of a copper wire and an iron wire fused together at the ends so as to form a circuit with two junctions. One of these, C, is at 0°C, and the other, H, is at 200°C. In Fig. 9 are plotted the thermoelectric powers of iron and copper, each with respect to lead.[2] Then according to the beautiful interpretation given, for example, by Starling,[3] and credited to P. G. Tait, we can take from this figure the four separate emf's in the copper-iron circuit, giving us the numbers of microvolts shown with their directions in Fig. 10. For since,

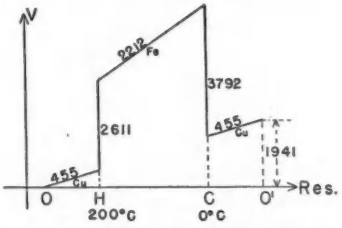


FIG. 11. Graph of potential against resistance in an open circuit copper-iron couple. The circuit is presumed to be broken at the midpoint of the copper wire. Potentials are quoted in microvolts.

as Starling and others[4] show, the Peltier emf π is given by Tde/dt, so that $\pi_1 = T_1(de/dt \text{ at } T_1)$ and $\pi_2 = T_2(de/dt \text{ at } T_2)$, we have $e = \text{Area } A_1A_2C_2C_1 = \text{Area } DA_2C_2J + \text{Area } JC_2C_1K + \text{Area } EA_1A_2D - \text{Area } EA_1C_1K = \pi_2 + e_T(\text{Fe}) + e_T(\text{Cu}) - \pi_1 = 2611 + 2212 + 910 - 3792 = 1941$ μv, clockwise in Fig. 10. The resulting current is accordingly clockwise in this diagram, or from copper to iron at the hot junction, as it is commonly said to be.

[2] *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio, 1945), twenty-ninth edition. In later editions these figures have not been changed.

[3] S. G. Starling, *Electricity and Magnetism* (Longmans, Green and Company, London, England, 1937), sixth edition, Chapter VII, p. 202. Essentially the same treatment is given in other editions.

[4] Leonard B. Loeb, *Fundamentals of Electricity and Magnetism* (John Wiley and Sons, Inc., New York, 1947), third edition, Chapter XII, p. 154. R. W. Hutchinson, *Advanced Textbook of Magnetism and Electricity* (W. B. Clive, London, England, 1917), Vol. II, Chapter XV, p. 169.

It appears, from Fig. 9, that if T_2 were increased to about 340°C, the two power lines would meet at the neutral temperature, giving maximum emf for these metals. For here $de/dt = 0$, and beyond this temperature e would be less. At this point, the area DA_2C_2J would become zero, corresponding to the fact that $\pi_2 = \pi_n = T_n(de/dt \text{ at } T_n) = 0$.

Since this paper was completed in first form, the author has been privileged to see proof of the article by E. T. Samuelson,[5] where a diagram like that of Fig. 9 is given. This figure is retained here to account for the numbers that appear in Fig. 10, which are to be used in what follows.

We are now able to construct the potential-resistance diagram for the copper-iron couple. For convenience we take the two resistances to be the same. We assume also that the temperature gradient is the same all along each wire, though not necessarily the same in iron as in copper. This would not be the case in a long wire with ends at 0°C and 200°C but could be approximately true in a short wire well insulated. For then most of the heat entering the wire at one end would leave it at the other end. The advantage of a uniform gradient is that then the Thomson emf per unit length and per unit resistance will be the same through the iron wire and through the copper wire, though different in the two. Then the potential along either wire, when plotted against resistance, gives a straight line, since the Thomson coefficient σ is defined by the equation $\Delta e = \sigma \Delta t$. Since the fall due to r also (the ir drop) is proportional to resistance, our graph will consist of straight lines. The potentials with the copper wire broken at the middle are shown in Fig. 11 so that a difference of 1941 μv appears at the break. In Fig. 12, however, we have no break, and when we have traversed the circuit we arrive at the potential of the starting point. In Fig. 11, obviously, 2212μv is the Thomson emf in the iron wire, and 455μv is half of the 910μv which is that in the copper wire. To find the corresponding rises with the circuit closed, we note that 1941μv must be lost on account of resistance, half of this amount in each wire, since we have taken the two resistances to be equal.

[5] E. T. Samuelson, Am. J. Phys. 18, 557 (1950).

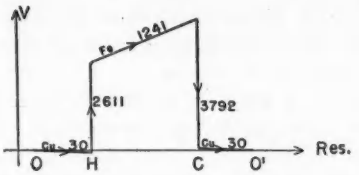


FIG. 12. Graph of potential against resistance in a closed circuit copper-iron couple. For simplicity, the resistance of the iron wire and that of the copper wire are assumed to be the same. Potentials are indicated in microvolts.

There remains a rise of 1241μv in the iron wire and falls of 30μv in each half of the copper wire. As Fig. 12 shows, there is almost no potential difference between the ends of the copper wire. The Thomson emf is almost large enough to cancel the ir drop. In the iron wire it is more than enough.

In order to calculate the efficiency of the thermal circuit as a thermodynamic device, it is of great advantage to have mechanical work done by the current. It might not be easy to construct, but we may think of the copper wire as constituting the armature of a motor, and of the resistances of the wires as so small that the ir products are negligible compared to 1941μv. Then this 1941μv will be the counter-emf of the motor that has a separately excited field. The essentials of this situation appear in Fig. 13. We noted in Fig. 7 that "The terminal potential of a motor is greater than the emf developed," and that is true of the common dc motor, but it is not the case in the thermal circuit. The reason is that here the ir drop is negligible, and moreover, the Thomson emf in the copper wire (910μv)

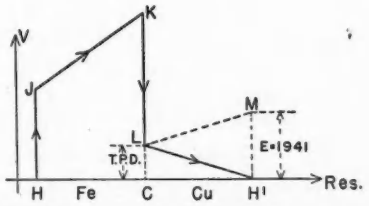


FIG. 13. Graph of potential against resistance for a closed copper-iron circuit of which the copper part is the armature of an ideal motor so that the mechanical work done by the current could be measured. Because the Thomson emf exists in the copper wire, the counter-emf of the motor (1941 microvolts) must be greater than the terminal potential difference (TPD). The hot and cold junctions H and C are taken, as before, at 200° and 0°C, respectively.

accounts for almost half of the counter-emf, so that the motor develops more volts than there are between its terminals.

It is worth while to construct the potential-resistance diagram of the copper-iron couple as it would have to be if there were no Thomson effects. In case the copper wire again forms the winding of an armature, and the *ir* terms are negligible, we have Fig. 14 instead of Fig. 13. Here we are obliged to make the rise in potential at *H* greater than the fall at *C*, or to say that the iron is at higher potential than the copper at each junction and more so at the hot junction. In no other way can we obtain a fall equal to a generated emf in the copper wire. But we see from Fig. 13 that actually the electricity, considered as positive, goes down in potential at the cold junction further than it rises at the hot junction. Accordingly, more energy is developed

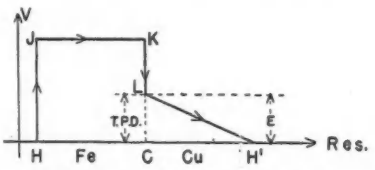


FIG. 14. This graph is a repetition of Fig. 13, except that Thomson emf's are presumed nonexistent. Now the counter-emf of the ideal motor is exactly equal to the terminal potential difference.

as heat at the cold junction of the copper-iron thermopile (with the kinds of material and the temperatures taken here) than is absorbed as heat and converted into electrical energy at the hot junction. Indeed, if the two Peltier emf's were as shown in Fig. 13 and there were no Thomson emf's, we should have a self-acting device doing mechanical work and also generating more heat than it absorbed. A glance at Figs. 13 and 14 shows how wrong it would be to draw any conclusions about the Peltier emf's without considering the Thomson effects.

It is certainly unnecessary to work out the details of a thermocouple of other materials, but we may add here that if we had used what is listed simply as iron in the Handbook,[2] instead of the transformer iron actually used, we should have found the current to be from iron to copper at the hot junction, and, of course, the Peltier emf's quite different. From iron to copper at the hot junction is, in fact, the direction of the current given by Sears.[6]

We cannot close this part of our review without studying the argument by which Sir William Thomson predicted the existence of an emf where there is a temperature gradient in a homogeneous conductor, which we now call the Thomson emf. It is an argument that we shall not be able to follow.

Thomson suspected the assumption that the only emf's in the circuit were the Peltier emf's (at the junctions), and he predicted the existence of a potential difference between points at different temperatures within a single conductor. However, so far as we can tell, he did not use the simple thermodynamic argument given, for example, by Starling, which shows that, in the absence of Thomson effects, the emf of a couple will be proportional to the difference of temperature between the junctions, this conclusion being contradicted by observations. We have in mind that if one junction of the copper-iron couple is kept at T_1, taken here as 0°C, and the temperature T_2 of the other is increased from T_1, then e passes through a maximum value when $T_2 = T_n$, which is called the neutral temperature. Thomson attributes the demonstration of this maximum property to Cumming,[7] and then says: "It is clear that, at exactly that temperature of the hot junction for which in any such case the electro-motive force is a maximum, the two metals must be thermo-electrically neutral to one another, and must present reverse thermo-electric relations for temperatures below and above this point. Hence the *thermal effect depending on the direction of a current crossing the junction of two such metals* must be for temperatures above, the reverse of what it is for temperatures below, the neutral point, and must vanish when the metals are exactly at this temperature." The italics are used by Thomson.

Now if we have a copper-iron couple with one junction at T_1 which is kept constant at 0°C, for example, and the other junction at T_2 which

[6] F. W. Sears, *Principles of Physics* (Addison-Wesley Press, Inc., Cambridge, Massachusetts, 1947), Vol. II, *Electricity and Magnetism*, p. 167.

[7] Sir William Thomson, *Mathematical and Physical Papers* (University Press, Cambridge, England, 1884), Vol. II, p. 195. This paper is the Bakerian Lecture, "On the Electro-dynamic Qualities of Metals," published also in the Trans. Roy. Soc. (London), February, 1856.

is gradually increased from zero, the total electromotive force of the couple increases to a maximum when $T_2=T_n$. If π_2 is $f(T_2)$, where f represents some function not specified, then the words quoted certainly means that if $f(T_2)-\pi_1$ shows a maximum, say, at $T_2=T_n$, then $f(T_n)$ must be zero. Far from being clear, this is obviously not the case. For $f(T_2)-\pi_1$ will have a maximum value where $f(T_2)$ is a maximum, whether that maximum $f(T_n)$ is zero or not. It may seem that we must have misunderstood the language of Thomson, but after a few lines about Peltier he continues: "If, therefore, a circuit of two metals have one junction kept at the neutral point, and the other at some lower temperature, the current excited will cause the evolution of heat at the cold junction, but neither absorption nor evolution of heat at the hot junction; and in the rest of the circuit there will be effects either purely thermal, or thermal and mechanical or chemical, according to the nature of the resistance against which the electro-motive force is allowed to work. The source from which the electro-motive force derived its energy to produce these effects cannot be at the hot junction, where heat is neither absorbed nor evolved, nor at the cold junction, where heat is evolved, nor of course in any uniformly heated part in either metal, through all of which, provided the metal has no thermo-electric crystalline characteristic, there can be nothing but a frictional evolution of heat; that is, it is nowhere but in those portions of the circuit where the temperature varies between that of the cold and that of the hot junction. In those portions, therefore, there must be as much heat absorbed, in virtue of the current, as is equivalent to the aggregate mechanical value of the heat evolved at the cold junction, and all the effects, thermal, mechanical, and chemical, produced in the rest of the circuit."

These words confirm our understanding of the argument. However, the conclusion is correct. For it follows from the application of the second law, as we have seen, that for the Peltier emf at T_2 we have $\pi_2=T_2(de/dT$ at $T_2)$. Since obviously at T_n, where e is a maximum, $de/dT=0$, $\pi_n=0$. Then, of course, the conclusion about emf and temperature gradient necessarily follows.

We are not the first, apparently, to find difficulty with Thomson's argument. Maxwell[8] and Livens,[9] for example, accept his statement to the effect that π_n is zero without commenting on his argument. Watson and Burbury[10] say simply that the Peltier effect changes sign at the neutral temperature. Hutchinson is more helpful here because he says (Ref. 4, p. 178): "From theoretical considerations Sir William Thomson was led to assume that the Peltier effect is zero when the thermo-electric power is zero." That is, of course, at T_n, where $de/dT=0$. Carhart[11] also says that "· · · Lord Kelvin assumed that the Peltier effect becomes zero at the neutral temperature." However, though we are unable to follow Thomson's argument, we are not obliged to call his conclusion an assumption. For he knew that at the neutral temperature the order of copper and iron is reversed in the thermo-electric series; that is, iron is at higher potential than copper at temperatures below T_n, and at lower potential above T_n, if two pieces of these metals are in contact, there being no complete circuit and no current whatever. For he says (reference 7, Vol. I, p. 321) "· · · the experiments made by Becquerel lead to the conclusion that at a certain high temperature iron and copper change their places in the thermo-electric series (a conclusion which the author has experimentally verified) and if this temperature be denoted by S we must consequently have $\Theta^S=0$." Here Θ^S has the meaning of our π_2. We must conclude accordingly that in his lecture Thomson did not set forth his real reasons for the statement that $\pi_n=0$, but gave what might have seemed at the moment to be reason enough, since he was sure of the fact. He would have saved one—perhaps the last—of his readers a good deal of distress if he had cited Becquerel and the phenomenon of reversal in the Bakerian lecture.

---

[8] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Clarendon Press, Oxford, 1873), Vol. I, p. 304.

[9] G. H. Livens, *The Theory of Electricity* (University Press, Cambridge, 1918), p. 307.

[10] H. W. Watson and S. H. Burbury, *The Mathematical Theory of Electricity and Magnetism* (The Clarendon Press, Oxford, 1885), p. 245.

[11] H. S. Carhart, *Physics for University Students* (Allyn and Bacon, Boston, 1904), Part II, p. 301.

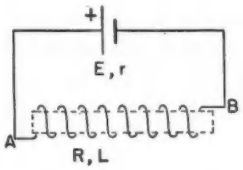


FIG. 15. A simple inductive circuit; the self-inductance L is independent of the current.

VI. SELF-INDUCTANCE. THE ENERGY OF THE MAGNETIC FIELD

What we have called electrical energy appears in its simplest form when two particles with positive charges Q and q are brought nearly together from a great distance. Mechanical energy is expended in this process, and an equal amount appears in the electrical form. The required work is given, as in any case, by an area, if the required force is plotted against the distance, and the amount is given by $V_A q$, if q is brought to the point A, V_A being the potential at A due to Q. We have seen that the charged condenser is the seat of electrical energy in amount $\frac{1}{2}VQ$, and if there is a uniform field between the plates it is easy to prove that the energy per unit volume of the medium, of dielectric constant k, is $kE^2/8\pi$ ergs cm^{-3}. The analogous expression for the energy per unit volume in a magnetic field is $\mu H^2/8\pi$. When a magnetic field is being established by building up a current in an inductive circuit, there is a conversion of electrical into magnetic energy, and when the current and the accompanying magnetic field decrease, the magnetic energy that disappears is transformed into electrical energy again. We can apply principles A and B (p. 263) to these changes advantageously.

Let us have, for example, a storage cell of emf E and resistance r connected, as in Fig. 15, to a spool of wire of resistance R and self-inductance L. There are no ferromagnetic materials about, so that L is constant, depending only upon the constant μ, which is probably close to unity, and the geometry of the circuit. If a quantity of electricity q crosses any (and every) section of the circuit, the cell loses the amount Eq of its store of chemical energy. While the current increases after the switch is closed, the magnetic field through and around the coil is being built up, and the energy of this field can appear only at the expense of electrical energy. According to principle B, electricity, considered as positive, is going from higher to lower potential while the current i is increasing, on account of the inductance, and of course, is going from higher to lower potential on account of resistance while there is a current, increasing or not. Then if the current i is directed from end A to end B of the coil, these facts and the definition of L indicate that $V_A - V_B = Ri + Ldi/dt$. Both R and L contribute to setting A above B in potential. If i, still directed from A toward B, is becoming less, the magnetic field is decreasing also, its energy is being returned to the electrical form, and therefore positive electricity must be moving from a point of lower to a point of higher potential on this account. This tends to make V_A less than V_B. This is provided for in the above equation, since now di/dt is negative. One may think of $V_A - V_B$ as the potential difference necessary to maintain the current and make it change, or as the potential difference appearing between A and B on account of a current maintained and compelled to change by an emf elsewhere in the circuit.

The foregoing argument makes it unnecessary to talk about an "induced" emf and an opposite "inductive" emf, and avoids all uncertainty about the sign of the Ldi/dt term. A situation exactly similar exists in mechanics, where an object is accelerated. If the frictional resistance is proportional to the velocity, we have for the required force (the force exerted by the agent that produces the motion) the equation $F = Kv + Mdv/dt$, where v and M denote velocity and mass, respectively, and K is a constant. Here it is not necessary to confuse the student by introducing the kinetic reaction $-Mdv/dt$ directed against the agent that makes the body move.

An expression for the energy represented by the magnetic field is easily obtained in the case

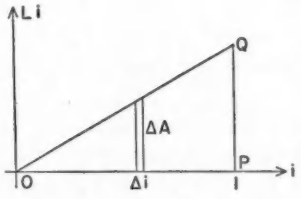


FIG. 16. In the inductive circuit of Fig. 15, the product Li is plotted as a function of i. On such a graph an area denotes energy. Here the energy required to establish the current I is seen to be $\frac{1}{2}LI^2$.

of no iron and constant L. It is only necessary to plot Li as function of i, as in Fig. 16, obtaining a straight line through the origin. Here, the variable current increases up to some value I in any manner whatever. Then any small change in current is the segment Δi, and ΔA, the area of the corresponding strip, is approximately $Li\Delta i$. Introducing Δt, the interval required to make the change, gives $\Delta A = L(\Delta i/\Delta t)i\Delta t$, approximately. The exact area A inclosed is the limit, as Δi and ΔA approach zero, of the sum of such terms as ΔA, and this limit is what is meant by the integral, so that

$$A = \int_{0=i}^{i=I} Li(di/dt)dt.$$

But here Ldi/dt is the potential difference necessary to produce the changing, and $idt = dq$, the element of charge crossing any section of the circuit; hence,

$$A = \int_0^Q edq = W,$$

the energy that must be supplied as i changes from zero to I in order to construct the magnetic field, Hence, the inclosed area is the energy represented by the field. Therefore,

$$W = \tfrac{1}{2}LI \times I = \tfrac{1}{2}LI^2.$$

Here W is in joules if L is in henries and I in amperes.

Exactly the same thing may be done in mechanics if we plot Mv as a function of the variable v between $v=0$ and $v=V$, and we obtain the kinetic energy W in the form $\frac{1}{2}MV^2$, no matter

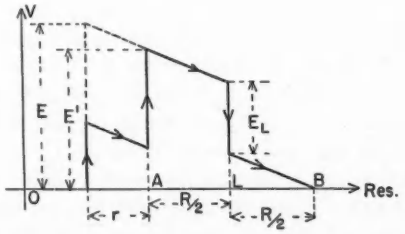


FIG. 17. Graph of potential against resistance at one particular instant when current is increasing in the inductive circuit of Fig. 15. The inductance L is assumed to be at the midpoint of the external resistance R. When the current has reached its final or steady value, the graph has evolved into the form of Fig. 1(B).

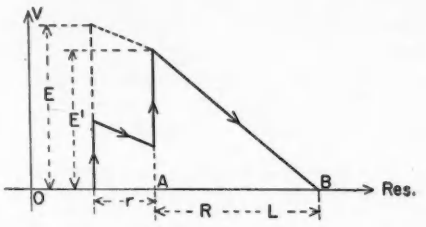


FIG. 18. This graph is a repetition of Fig. 17, except that the inductance L is uniformly distributed along the resistance R. Here, the inclined lines have different slopes until the current reaches its final or steady value.

how v may vary, and no matter how the force that generates the final velocity may vary.

In the case of the dc motor, we deduced the the equation $TPD = i_A r_A + E$, where E, commonly called the counter-emf, is the emf developed in the armature. Now we have, for the inductive coil, $V_A - V_B = Ri + Ldi/dt$. Accordingly, we call Ldi/dt the counter-emf of self-induction. At any rate, the counter-emf here is Ldi/dt, not $-Ldi/dt$, if counter-emf is to have the same meaning in the coil and in the motor.

The potential-resistance diagrams for the inductive coil are now easily constructed. It is necessary, however, to locate the inductance with respect to the resistance, and there will be two cases. In the first, L is in a coil of resistance zero or negligible at the middle of R, so that $R/2$ is on each side of L outside the generator, which we take to be a cell. Figure 17 shows the situation at any time after the circuit is closed but before the variable current i has reached its final or steady value I. The slope of the inclined lines is the instantaneous current i; E is the emf of the cell, and E' is the terminal potential difference, so that

$$E' = E - ir = iR + E_L = iR + Ldi/dt.$$

When i has attained to its final and constant value, I conditions are represented by Fig. 1(B).

However, if the circuit external to the cell consists of a coil of wire with the inductance L distributed uniformly with respect to the resistance R, we have the circumstances of Fig. 18. Here the inclined lines have different slopes.

We now suppose that i has attained (nearly enough) to its final or constant value I and then the external circuit (with R and L) is suddenly shorted. The consequences to the cell need not

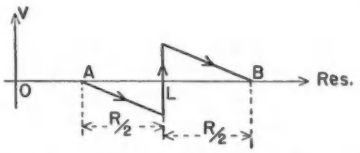


FIG. 19. Graph of potential against resistance at one particular instant, while the current is decreasing for an inductive circuit in which the inductive-resistive part was shorted when current was flowing. The inductance L is assumed to be localized at the midpoint of the resistance R.

concern us, or we can suppose that the cell is disconnected so soon after the short-circuit that no damage is done. Then the current begins to diminish at once, and since the magnetic field also decreases, the energy of the field becomes less as it is converted into the electrical form. With the inductance localized at the middle of R, positive electricity goes from lower to higher potential in L, as indicated in Fig. 19. Here $E'=iR+Ldi/dt=0$, and di/dt is negative.

The case of distributed inductance is shown in Fig. 20. The circuit is, for example, a complete ring solenoid, rather closely wound and without iron, and shorted when the current is not zero. Then, as the current falls to zero, there are no potential differences whatever in the conductor, for the rise due to inductance balances the fall due to resistance in each separate turn, even in each separate micron of the wire. Figure 20 represents the situation for every value of the current from I to zero. This case is quite analogous to that of a stone thrown across level ice. As the motion continues for an appreciable time in spite of friction, ma or mdv/dt balances the resisting force, or supplies the force necessary on account of friction, so that $0=F_f+mdv/dt$.

A short-circuited solenoid is not the only conductor in which we may have a current without any potential difference whatever. Let there be an alternating current of 5 amperes, for example, in a helical coil with an open core made of soft iron wires. If a uniform ring made of copper, of cross-sectional area 1 cm², for example, is held over or outside the helix, a very large current will be induced in the ring, indicated by the fact that the ring becomes uncomfortably hot in a few seconds, though its resistance is very small, perhaps 4×10^{-5} ohm. But no point in the ring will be at higher potential than any other

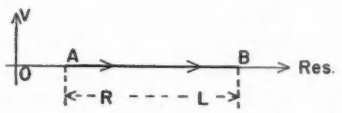


FIG. 20. A repetition of Fig. 19 for the case in which the inductance L is uniformly distributed along the resistance R. Between any two points of the conductor there are two differences of potential, one of which opposes and exactly balances the other.

point. No point is electrically distinguishable from any other. The open-core transformer and metal ring are commonly used to show electromagnetic repulsion. An important feature of the experiment, not to be ignored by the operator, is that the current in the helix is much higher when the ring is in place around it, very likely 20 amperes instead of the original 5. Students may be asked to explain the increase.

We may close with a remark about Lenz's law. It is remarkable that there are may different statements attributed to Lenz in general physics books, and not one of them have I found to be a translation, not even a free translation, of what Lenz[12] actually wrote. The following is a literal translation.

"If a metallic conductor moves in the neighborhood of a current or a magnet, a current is generated having a direction such that it would produce motion of the wire opposite to that actually given to it, supposing that the wire was movable only in the direction of its motion and opposite."

Since the law thus stated does not specifically apply to the emf of self-induction, nor to mutual induction, which are at least as important as emf's due to the motion of a conductor, it is not strange that recent writers have generalized the statement by Lenz, but they ought not to attribute to him things that he did not say. In the case of self-inductance it is commonly said that the emf is so directed as to oppose the change of current that produces it. Principles A and B (p. 263) may be used equally well to determine the direction, and the argument by the transformation of energy seems more fundamental and just as easy.

It may be remarked in conclusion that prin-

[12] E. Lenz, Ann. Physik Chem. 31, 483 (1834).

ciples A and B are not limited in application to the situations taken up in this paper. In particular, the motion of an electron toward or away from an atomic nucleus represents a transformation between mechanical and electrical energy, or, when the electron changes from one stable state to another, between electrical energy, kinetic energy of motion and light. However, no electrical resistance is involved in this motion of the electron, and the process would have to be forced unnaturally into the potential-resistance diagram.

The Orientation of Paramagnetic and Diamagnetic Rods in Magnetic Fields

ARTHUR R. LAUFER
University of Missouri, Columbia, Missouri
(Received January 29, 1951)

It is commonly believed that a paramagnetic rod will align itself with an impressed magnetic field and that a diamagnetic rod will set its long dimension across the field. The present paper shows that this conception is not generally valid. In a strictly uniform field both paramagnetic and diamagnetic rods will set themselves parallel to the field. In a nonuniform field the orientation of each of these rods will depend upon the particular configuration of the field.

FARADAY designated two classes of substances as paramagnetic (Greek *para*, beside) and diamagnetic (Greek *dia*, through) when he found that slender rods of these materials oriented themselves parallel and perpendicular, respectively, to the direction of the field between the poles of his electromagnet. His designations are undoubtedly responsible for the prevalent but, as it will be shown, erroneous view that it is an intrinsic property of paramagnetic and diamagnetic rods that they will set themselves parallel and perpendicular, respectively, to the lines of force of a magnetic field. That this misconception is furthered by most textbooks is clearly evident in the following two typical quotations:[1]

1. "The existence of the magnetic properties of a substance can be demonstrated by supporting a small rod-shaped specimen at its center of gravity by a fine thread and placing it in the magnetic field of a powerful electromagnet. Everyone knows that if the specimen is of iron or one of the *ferromagnetic* substances it will align itself in the direction of the magnetic field. Not as familiar is the fact that any substance whatever will be influenced by the field, although to an extent which is extremely small compared with a substance like iron. Some substances will, like iron, set themselves with their long dimension parallel to the field, while others will come to rest with their long dimension perpendicular to the field. The first type are called *paramagnetic*, the second, *diamagnetic*."

2. "If we place oblong pieces of these substances in a *uniform field* the iron-like substances will set themselves parallel to the field so that as many lines of force as possible run through the body. The others will set themselves at right angles to the field trying to avoid the passage of magnetic lines. The bodies which ···*set themselves parallel to the uniform field are called paramagnetic bodies.*··· Those bodies which···*set themselves perpendicular to the uniform field are called diamagnetic bodies.*"

[1] The italics are those of the authors of these quotations.

In point of fact, these two classes of substances should be specified in terms of the signs of their magnetic susceptibilities, which in turn appear to be associated with the atomic and electronic properties of matter. By virtue of its spin and orbital motion each electron in an atom contributes a magnetic moment to the atom, the spin contribution being of greater importance. If the spin of each electron in the atom is balanced by that of another electron with opposite spin, the permanent magnetic moment of the atom as a whole vanishes. If the electron spins of the atom are not balanced, the atom will exhibit a nonvanishing permanent magnetic moment. The incomplete $4f$ subshell of the rare earths, for example, is the cause of their strong paramagnetism. An external magnetic field applied to a substance with permanent atomic moments serves to align the moments in the direction of the field, resulting in magnetization of the substance in the same direction as the im-

pressed field. On the other hand, the atoms of all substances, when subjected to an external field and whether or not they possess permanent moments, exhibit small induced magnetic moments which have a direction opposite to that of the impressed field.

If the atoms of a substance possess permanent moments, the alignment effect far overshadows the induced effect and the net magnetization of the substance is in the same direction as the impressed field. But if the permanent moments vanish the magnetization, now entirely ascribable to the induced moments, is opposite to the direction of the field. In both cases the magnetization, $\mathbf{M}$, is given in gaussian units by $\mathbf{M}=\chi\mathbf{H}$, where $\mathbf{H}$ is the magnetic field intensity in the substance and χ is the magnetic susceptibility. With the exception of certain paramagnetic materials at low temperatures and of those diamagnetic materials which exhibit the de Haas-von Alphen effect, the susceptibility is usually assumed to be independent of the field intensity for paramagnetic and diamagnetic substances. In general, χ has the form of a dyadic; but if the medium is magnetically isotropic, χ is a scalar function of the coordinates. If the medium is also homogeneous, χ reduces to a scalar constant. For the class of substances known as paramagnetic χ has a small positive value, for example, 29×10^{-6} for platinum. For the class known as diamagnetic χ has a small negative value, for example, -14×10^{-6} for bismuth. The magnetic susceptibility is related to the magnetic permeability, μ, by the equation $\mu=1+4\pi\chi$.

The orientation of paramagnetic and diamagnetic substances in magnetic fields is determined by the mechanical forces and torques acting on the substances. It follows[2] from Maxwell's equations that the mechanical force on a unit volume of a substance in a magnetic field is

$$\mathbf{f}=(1/c)\mathbf{i}\times(\mathbf{B}-4\pi\mathbf{M}_0)-(1/8\pi)H^2\nabla\mu+\mathbf{H}\nabla\cdot(-\mathbf{M}_0), \quad (1)$$

where $\mathbf{i}$ is the electric current traversing the substance, $\mathbf{B}$ is the magnetic induction, and $\mathbf{M}_0$ is the residual magnetization in the absence of an impressed field. Paramagnetic and diamagnetic substances show no residual magnetization, and if no current traverses the substance, Eq. (1) reduces to $\mathbf{f}=-(1/8\pi)H^2\nabla\mu$. To express the force density in terms of the susceptibility, μ may be replaced by $\mu-1$ to yield

$$\mathbf{f}=-\tfrac{1}{2}H^2\nabla\chi. \quad (2)$$

The total force on a rigid body of volume V in a vacuum is then

$$\mathbf{F}=-\frac{1}{2}\int_V H^2\nabla\chi dv, \quad (3)$$

it being necessary to postulate the usual transition layers at the surface of the body through which χ varies continuously, although rapidly, from $\chi=\chi$ to its value $\chi=0$ in the surrounding vacuum. Writing

$$H^2\nabla\chi=\nabla(H^2\chi)-\chi\nabla H^2, \quad (4)$$

Eq. (3) becomes

$$\mathbf{F}=-\frac{1}{2}\int_V\nabla(H^2\chi)dv+\frac{1}{2}\int_V\chi\nabla H^2dv. \quad (5)$$

But $\int_V\nabla(H^2\chi)dv=\int_S H^2\chi d\mathbf{s}$, where S is the surface bounding the volume V. If this surface is taken to be immediately adjacent to the surface of the body but in the surrounding vacuum where $\chi=0$, the surface integral vanishes, leaving

$$\mathbf{F}=\frac{1}{2}\int_V\chi\nabla H^2dv. \quad (6)$$

The total torque on the rigid body is $\mathbf{L}=\int_V\mathbf{r}\times\mathbf{f}dv$, where $\mathbf{r}$ is the position vector of the element of volume dv and $\mathbf{f}$ is given by Eq. (2). Therefore, $\mathbf{L}=-\frac{1}{2}\int_V\mathbf{r}\times H^2\nabla\chi dv$, which by Eq. (4) becomes

$$\mathbf{L}=-\frac{1}{2}\int_V\mathbf{r}\times\nabla(H^2\chi)dv+\frac{1}{2}\int_V\mathbf{r}\times\chi\nabla H^2dv. \quad (7)$$

Now, $\mathbf{r}\times\nabla(H^2\chi)=H^2\chi\nabla\times\mathbf{r}-\nabla\times(H^2\chi\mathbf{r})$, and as $\nabla\times\mathbf{r}=0$, the first term in Eq. (7) becomes $\frac{1}{2}\int_V\nabla\times(H^2\chi\mathbf{r})dv$. But $\int_V\nabla\times(H^2\chi\mathbf{r})dv=-\int_S(H^2\chi\mathbf{r})\times d\mathbf{s}$, where $d\mathbf{s}$ is an element of the surface S surrounding the volume V. Again, on this surface $\chi=0$ everywhere and the surface integral vanishes, leaving

$$\mathbf{L}=\frac{1}{2}\int_V\mathbf{r}\times\chi\nabla H^2dv. \quad (8)$$

[2] M. Abraham and R. Becker, *The Classical Theory of Electricity and Magnetism* (Blackie & Son, London, 1942).

Evidently, by Eqs. (6) and (8),

$$\mathbf{f}' = \tfrac{1}{2}\chi \nabla H^2 \qquad (9)$$

may be regarded as the mechanical force of magnetic origin on each unit of volume of a rigid body in a vacuum. In this expression for the force density, $\mathbf{H}$ is the actual magnetic intensity inside the rigid body.

Let it first be supposed that the body is in a uniform external field, $\mathbf{H}_0$. As χ is very small, it might appear that to a first approximation the small induced magnetization alters the field within the body only negligibly, and that consequently the field $\mathbf{H}$ in the entire region occupied by the body is effectively the same as if the body were not there. Therefore, $\mathbf{H}=\mathbf{H}_0$, ∇H^2 vanishes, and by Eqs. (6) and (8) both the force and torque on the body vanish. Actually, of course, this approximation and these results are not valid. The alteration of the field in the region occupied by the body cannot be neglected. As a matter of fact, a resultant force or torque exists only to the extent to which the introduction of the body *does* alter the uniform field so that ∇H^2 does not vanish.

The presence of the body alters the original magnetic intensity in the region it occupies by virtue of the magnetization induced in it. Associated with this magnetization is a new magnetizing force known as the demagnetizing field, $\mathbf{H}'$, which must be subtracted from the impressed intensity, $\mathbf{H}_0$, to find the true intensity, $\mathbf{H}$, inside the body. The demagnetizing field is usually specified by a demagnetizing factor N defined by $\mathbf{H}'=N\mathbf{M}$. Hence, the true field inside the body is $\mathbf{H}=\mathbf{H}_0-\mathbf{H}'=\mathbf{H}_0-N\mathbf{M}$. The demagnetizing factor, which is a function of the geometric form of the body, is in general not easily calculable. However, if the form of the body is such that the magnetization within it is uniform, the demagnetizing factor can be calculated without undue difficulty.

The magnetization created by a uniform impressed field within a homogeneous isotropic specimen in the form of a cylindrical rod is not uniform throughout the rod. It is only when the specimen is in the form of an ellipsoid that the magnetization within it is uniform. Hence, it is convenient for the purpose of quantitative analysis to treat the cylindrical rod as a long prolate ellipsoid of revolution. If then the impressed uniform field is in the direction of the major axis of a prolate ellipsoid of revolution of eccentricity e, the demagnetizing factor has been shown[3] to be

$$N_1 = 4\pi\frac{1-e^2}{e^2}\left(\frac{1}{2e}\log\frac{1+e}{1-e}-1\right),$$

whereas if the impressed field has the direction of a minor axis, the demagnetizing factor is

$$N_2 = 2\pi\left(\frac{1}{e^2}-\frac{1-e^2}{2e^3}\log\frac{1+e}{1-e}\right).$$

The sum of the demagnetizing factors along the three axes of an ellipsoid has the value 4π, so for the prolate ellipsoid, $N_1+2N_2=4\pi$. It can also be verified that $N_1<N_2$. Values of the demagnetizing factors have been calculated and tabulated for various values of the ratio of length to diameter for the prolate ellipsoid of revolution, and have been measured and tabulated for the same values of this ratio for a cylindrical rod.[4] The satisfactory agreement justifies the treatment of a cylindrical rod as a prolate ellipsoid of revolution in the following analysis.

Consider now a prolate ellipsoid of revolution with its major axis in the XY plane and making an angle θ with an impressed uniform field, $\mathbf{H}_0$, which has the direction of the X axis. Within the body of the ellipsoid the field, $\mathbf{H}$, is uniform and ∇H^2 vanishes; but at the surface of the ellipsoid the intensity varies rapidly from its value inside to its value outside and ∇H^2 does not vanish. Consequently, the ellipsoid experiences a torque about the Z axis, the magnitude of which has been calculated by Maxwell.[3] In the notation employed above, and under the further assumption that the ellipsoid is magnetically homogeneous and isotropic, Maxwell's result may be written as

$$L_z = \chi^2 V H_0^2 (N_1-N_2)\sin 2\theta/2(1+N_1\chi)(1+N_2\chi),$$

where V is the volume of the ellipsoid. If χ is

[3] James Clerk Maxwell, *Treatise on Electricity and Magnetism* (Oxford University Press, 1892), third edition, Vol. 2.

[4] R. Becker, Z. Physik **87**, 547 (1934); R. M. Bozorth and D. M. Chapin, J. Appl. Phys. **13**, 320 (1942).

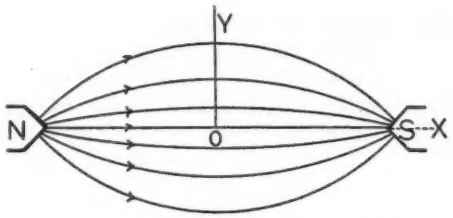


FIG. 1. Magnetic field between conical pole pieces.

very small, this equation reduces to

$$L_z = -(\chi^2/2)VH_0^2(N_2-N_1)\sin 2\theta. \quad (10)$$

Now, N_1 and N_2 are intrinsically positive, and $N_2>N_1$. Therefore, for $\theta<\pi/2$ the torque seeks to diminish the angle between the ellipsoid's major axis and the direction of the impressed field. The ellipsoid tends, evidently, to align its major axis with the impressed field no matter whether χ is positive or negative. Consequently, both paramagnetic and diamagnetic rods in a *uniform* field, if free to turn, will seek to set their long axes parallel to the field. This is in direct contradiction to the usual textbook statement that a diamagnetic body in a uniform field tends to set its longest dimension athwart the field. It will not do so if the field is strictly uniform.

However, this χ^2-proportional torque has probably never been observed. In a presumably uniform field the exceedingly small value of χ^2 for diamagnetic substances yields a resultant torque of such small magnitude that even a slight nonuniformity of the field will alter the orientation of the rod, the torque resulting from the *lack of uniformity* of the impressed field overpowering that given by Eq. (10). Furthermore, the electromagnets usually employed to observe these orientation effects yield strongly nonuniform fields. It is therefore necessary to investigate the orientation of a rod in a nonuniform magnetic field. The slight magnetization of a paramagnetic or diamagnetic rod in a strongly inhomogeneous impressed field, $\mathbf{H}_0$, does not appreciably change the configuration of the field. Equation (9) may therefore be written as

$$\mathbf{f}' = \tfrac{1}{2}\chi\nabla H_0^2. \quad (11)$$

Two familiar configurations of the field $\mathbf{H}_0$ are of particular interest. The first, as indicated in Fig. 1, is the type of field which is usually found between the conical poles of an electromagnet. The second, illustrated in Fig. 2, is the type produced by a current-carrying coil. These fields are not only symmetrical about the X axis but also have mirror-image symmetry with respect to the YZ plane through the origin, O. The general form of the lines of force in each case suggests that, at least in the neighborhood of the origin, each component of the field may be expanded as a power series in the coordinates. If the symmetry conditions are applied to these expansions, together with Maxwell's equations specifying that both the divergence and the curl of $\mathbf{H}$ must vanish in this region, the field is found to be given by

$$H_0^2 = A + Bx^2 - (B/2)(y^2+z^2), \quad (12)$$

where A and B are constants. If B is positive, the magnetic intensity and the line density increase as one moves along the X axis away from the origin in either direction, whereas the intensity and density decrease as one advances perpendicularly out from the X axis. This obviously corresponds to the configuration of Fig. 1. But if B is negative, the configuration of Fig. 2 results, the intensity and density decreasing in either direction along the X axis and increasing perpendicularly out from the X axis.

Consider a slender cylindrical rod, of length d and cross-sectional area s, lying in the XY plane with its center at the origin. The components of force per unit volume in this plane are, by Eqs. (11) and (12),

$$f_x' = \tfrac{1}{2}\chi\partial(H_0^2)/\partial x = \chi Bx,$$
$$f_y' = \tfrac{1}{2}\chi\partial(H_0^2)/\partial y = -\tfrac{1}{2}\chi By.$$

From Eqs. (8) and (9) it is evident that an element of the rod of length dr and volume sdr, at a distance $r=(x^2+y^2)^{\frac{1}{2}}$ along the rod from the

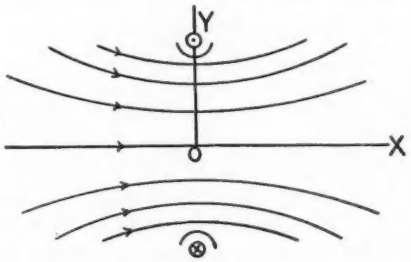


FIG. 2. Magnetic field of a current-carrying coil.

origin, experiences a torque about the Z axis given by

$$l_z=(xf_y'-yf_x')sdr.$$

If θ is the angle between the axis of the rod and the X axis, $x=r\cos\theta$, $y=r\sin\theta$, and the resultant torque on the entire rod is

$$L_z=\int_{-d/2}^{d/2}(-\tfrac{1}{2}\chi Br^2\sin\theta\cos\theta-\chi Br^2\sin\theta\cos\theta)sdr.$$

Multiplying and dividing by the constant density, ρ, of the rod, this becomes

$$L_z=-\chi(3B/4\rho)\sin2\theta\int_{-d/2}^{d/2}r^2\rho sdr, \qquad (13)$$

where the integral is merely the moment of inertia of the rod about the Z axis and is intrinsically positive.

It follows that in a field of the type of Fig. 1, for which B is positive the torque and the magnetic susceptibility have opposite signs for $\theta<\pi/2$. A paramagnetic rod ($\chi>0$) will seek to set itself longitudinally along the X axis, parallel to the lines of force, whereas a diamagnetic rod ($\chi<0$) will tend to set itself across the X axis and the lines of force. This field configuration is the one which commonly exists between the poles of an electromagnet; and, therefore, it generally *does* happen, as observed by Faraday, that a diamagnetic rod sets itself across the lines of force of the field. However, it should now be obvious that this orientation is not a property of the rod alone, but is merely a consequence of the particular configuration of the field. If B is negative, the torque will tend to set a paramagnetic rod *across* the lines of the field and a diamagnetic rod *along* the lines of the field, the orientations to be expected in fields of the type of Fig. 2.

The magnetic susceptibilities of paramagnetic and diamagnetic rods are so small in magnitude that intense fields are required in order that the orientation effects be observable. An iron-cored electromagnet is capable of producing the required intensity, but the usual plane or conical pole pieces can produce fields only of the type of Fig. 1. An air-cored current-carrying coil, which can produce fields of the type of Fig. 2, does not in general provide sufficient intensity. However, it should be possible to produce fields of this type and of the required intensity by furnishing an electromagnet with pole pieces of proper design.

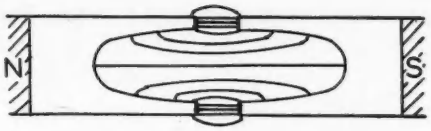


FIG. 3. Magnetic field between special pole pieces.

When lines of magnetic induction pass from a medium of permeability μ_1 into a second medium of permeability μ_2, the required continuity of the normal component of **B** and of the tangential component of **H** at the interface leads to the refraction of the lines according to the law

$$\tan\theta_2/\tan\theta_1=\mu_2/\mu_1,$$

where θ_1 and θ_2 are the angles made by the lines with the normal to the interface in the first and second mediums, respectively. If the first medium is iron and the second is air, μ_2/μ_1 is very small, θ_2 is very small for most values of θ_1, and the lines emerge from the iron in a direction nearly normal to the surface. Therefore, if two iron blocks containing properly shaped cavities, as indicated in Fig. 3, are attached to the pole pieces of an electromagnet, an intense field of the form of Fig. 2 will be produced in the region between the blocks. A diamagnetic rod in this magnetic field should align itself along the lines of force and a paramagnetic rod should set itself across the lines.

In conclusion, it may be peculiarly appropriate to quote the following passage from the Report[5] of the Coulomb's Law Committee of the American Association of Physics Teachers:

"While many statements in textbooks are undoubtedly wrong because the authors themselves have an incorrect understanding of the facts involved, others are at fault because the authors have chosen to introduce a general subject by the consideration of a special case and have not made clear to the student that the preliminary discussion is of limited scope. In some textbooks the general case is not mentioned."

[5] Am. J. Phys. 18, 1 (1950).

A Correction to the Treatment of Fresnel Diffraction*

C. L. ANDREWS
New York State College for Teachers, Albany, New York
(Received October 2, 1950)

Measurements of the diffraction pattern of microwaves near circular apertures reveal the quantitative correctness of Kirchhoff's obliquity factor. It has been customary in textbooks of physical optics to assume with Fresnel that the amplitude from zones an infinite distance from the center of the aperture is zero. However, according to Kirchhoff's equation, the amplitude from a zone as the zone number increases without bound is not zero but half the amplitude from the central zone. It is noted that a marginal correction in the textbooks will give results that agree not only with the observed behavior of light, but of microwave diffraction as well, up to the plane of the aperture. The more daring instructor may wish to make a more complete break with the past and explain the diffraction pattern on the basis of Thomas Young's theory. By Young's simple theory that the diffraction pattern of the aperture is an interference pattern between the incident wave and secondary waves from the edge of the aperture, the positions of maximum and minimum intensity may be predicted not only in the Fresnel region but also behind the aperture and in the plane of the aperture itself.

MEASUREMENTS of Fresnel diffraction patterns of circular apertures for hand-sized microwaves in and near apertures[1] suggest that the classical treatment of diffraction is in need of revision and that the study of diffraction patterns in the neighborhood of the diffracting apertures and objects will be a fruitful field for elementary research. However, we need not expect drastic revision in the elementary treatment of physical optics in the immediate future. Fresnel methods have served well in predicting positions of maximum and minimum intensities of diffraction patterns of a circular aperture near the axis of the aperture even up to the very aperture itself.

To be sure, microwave measurements reveal that Fresnel's assumption that the intensity and phase are constant over the aperture is false. The sharpest diffraction patterns and greatest intensity gradients are in the plane of the aperture. However, for large apertures the intensity in the plane of the aperture does not vary widely except for a couple dozen wavelengths from the center and a dozen wavelengths from the edge of the aperture.

The Fresnel zone method, like the Bohr picture of the atom, will be retained in spite of its deficiencies until some simpler, graphical, pedagogic picture can be found.

In the meantime the Fresnel treatment must receive one minor bit of patching. Most textbooks in physical optics[2–3] make use of Fresnel's scheme of grouping amplitudes of waves arriving at a point on the axis of the aperture from an infinite series of Fresnel half-wave zones in order to calculate the resultant amplitude. It is assumed that the amplitude of a wave from a zone approaches zero as the number of the zone increases without bound. Although this assumption is contrary to Kirchhoff's refinement of Fresnel's theory, it has been retained because it seemed to give the correct answer for intensities at points on the axis.

However, microwave measurements of intensities at points on the axis at a short distance from the aperture compared to the diameter of the aperture reveal the correctness of Kirchhoff's obliquity factor. Therefore, we should no longer ignore Kirchhoff's obliquity factor by assuming the amplitude of the wave from the nth zone to approach zero as n becomes large.

As an example, the necessary change in Jenkins and White's book[2] may be made by three marginal notes on three successive pages. On page

* The measurements and devices described in this paper were developed while the writer was acting as consultant to the General Electric Research Laboratory, Knolls I, Schenectady, New York.

[1] C. L. Andrews, J. Appl. Phys. 21, 761 (1950).

[2] F. A. Jenkins and H. E. White, *Fundamentals of Physical Optics* (McGraw-Hill Book Company, Inc., New York, 1937), first edition, p. 178.

[3] Houston, *A Treatise on Light* (Longmans, Green and Company, London, 1915), p. 156.

177, Kirchhoff's results for the amplitude of the wavelet reaching a point on the axis of the aperture from the nth zone is expressed by

$$R_n = \text{constant}(A_n/d_n)(1+\cos\theta), \quad (1)$$

where R_n is the amplitude from the nth zone, A_n is the area of the zone, d_n is the distance from the zone to the point on the axis, and θ is the angle between the direction of propagation of the unperturbed wave and a line from a point on the zone to the point on the axis. The ratio A_n/d_n is proven to be the same for all n's so that θ is the only variable. We may note that, as

$$n\to\infty, \quad \theta\to\pi/2,$$

and

$$\lim_{n\to\infty} R_n = \tfrac{1}{2}R_1, \quad (2)$$

where R_1 is the amplitude from the first zone. On page 178, by Fresnel's famous series method, it is shown that if n is odd,

$$\tfrac{1}{2}R_1+\tfrac{1}{2}R_n = R. \quad (3)$$

Here R is the resultant amplitude from all the zones. It is left for the student to prove that if n is even,

$$\tfrac{1}{2}R_1-\tfrac{1}{2}R_n = R. \quad (4)$$

If, instead of assuming, as Fresnel did, that $R_n=0$ when n is large, we substitute Kirchhoff's value of R_n from Eq. (2); then, if n is odd,

$$R = \tfrac{1}{2}R_1+\tfrac{1}{4}R_1 = \tfrac{3}{4}R_1, \quad (5)$$

and, if n is even,

$$R = \tfrac{1}{2}R_1-\tfrac{1}{4}R_1 = \tfrac{1}{4}R_1. \quad (6)$$

Finally, on page 179, Fig. 8*E*, in the amplitude diagram for half-period zones, the component amplitude should not approach zero as n becomes large. It should approach half the amplitude from the first zone in agreement with Eq. (1). There are at least two experimental reasons why, as the diameter of the aperture is increased, we do not observe the alternations in amplitude of light waves between $\frac{1}{4}R_1$ and $\frac{3}{4}R_1$ but instead observe the average of these two extremes:

(1) The radius of the aperture is never constant to a sufficiently small fraction of a wavelength.

(2) As pointed out by Crose and Boillet,[4] if the source has dimensions much greater than a wavelength, then the amplitude of wavelets from the outer zones will be zero.

With hand-sized microwaves each of these sources of error may be removed.

[4] F. Croze and P. Boillet, Compt. rend. **230**, 200–3 (1950).

KIRCHHOFF'S THEORY OF DIFFRACTION

It seems desirable to give increased attention to Kirchhoff's theory of diffraction in intermediate courses in physical optics for two reasons:

(1) The results obtained for the intensity along the axis of the circular aperture are in surprisingly good agreement with experimental measurements of light and microwaves.

(2) The physical meaning of the surface integral can be readily interpreted by the student, and, if no approximations are made, the solution of the integral has been shown[5] to be a simple case from the elementary integral tables.

The derivation of Kirchhoff's theorem from Green's theorem is found in books on optics[6] and electromagnetic theory.[7]

$$u_p = \frac{u_0}{4\pi}\int\int\left[\frac{2\pi}{\lambda}\frac{\epsilon^{-i\phi}}{\rho}(1+\cos\theta) - i\frac{\epsilon^{-i\phi}}{\rho^2}\cos\theta\right]d\sigma. \quad (7)$$

Here, u_p is the amplitude at point P resulting from the added effects from all points in the aperture, u_0 is the amplitude of the undisturbed incident wave at the aperture, $d\sigma$ is an element of area in the aperture, ρ is the distance from the element $d\sigma$ to P, ϕ is the phase lag of the component from $d\sigma$ behind the component from an element at the base of a perpendicular from the point P to the plane of the aperture, and θ is the angle between the incident beam and a line from $d\sigma$ to the point P.

In books on physical optics the second term is omitted, since $\rho\gg\lambda$ in all cases treated in light. However, in the study of microwaves, useful applications occur when ρ is of the same order as a wavelength. Note the direction factors $(1+\cos\theta)$ and $(\cos\theta)$ for each of the two terms.

[5] C. L. Andrews, Phys. Rev. **71**, 777 (1947).

[6] M. Born, *Optik* (Verlag. Julius Springer, Berlin, 1933), pp. 147–151.

[7] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), pp. 424–428, 460–464.

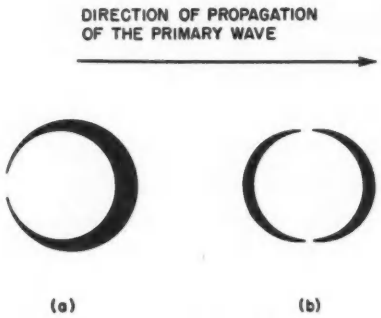


FIG. 1. Huygens' wavelets for the two terms in Kirchhoff's equation: (a) the wavelet whose direction factor is $(1+\cos\theta)$ and (b) the wavelet whose direction factor is $\cos\theta$.

Figures 1(a) and 1(b) indicate relative amplitudes of Huygens' wavelets in all directions for each of the two terms. If the study is restricted to determination of disturbances along the axis of the aperture, then it applies to vector as well as to scalar wave functions. The solution[5] yields

$$u_p/u_0=(\sin\beta+i\cos\beta)\times[(\beta+4\pi n)/(2\beta+4\pi n)]-i, \quad (8)$$

where β is the phase of the component from the outermost subzone of the aperture relative to that from the center of the aperture and n is the distance in wavelengths of P from the center of the aperture. Note that this is the equation of a spiral with center at $-i$ lying inside a circle of unit radius and approaching a circle of radius $\frac{1}{2}$ as asymptote. Kirchhoff's spiral differs from Fresnel's spiral, which approaches a point as asymptote.

To express the amplitude at P in terms of the diameter of the aperture, we may let d be the diameter. Then,

$$d/\lambda=[(\beta/\pi)^2+4n\beta/\pi]^{\frac{1}{2}}. \quad (9)$$

Figure 2 is a plot of the imaginary *versus* the real components of the amplitude for values of β from 0 to 7π for distances along the axis from the center of the aperture of 5.0 and 0.5 wavelengths. The broken circle is the asymptote approached by the spiral.

If n is large, as is usually so for most measurements of diffraction patterns of light, then even the first turns of the spiral are so tight that they are shown only schematically in the textbooks. For the small number of zones commonly treated in laboratories of physical optics, when an odd number of zones are uncovered, the amplitude is twice that of the unperturbed beam, and when an even number is uncovered, the amplitude is close to zero.

If microwaves are employed and measurements made at points a few wavelengths from the aperture, then the effect of the obliquity factor may be observed. The closing of the spiral can be measured.

For n's smaller than unity the spiral closes

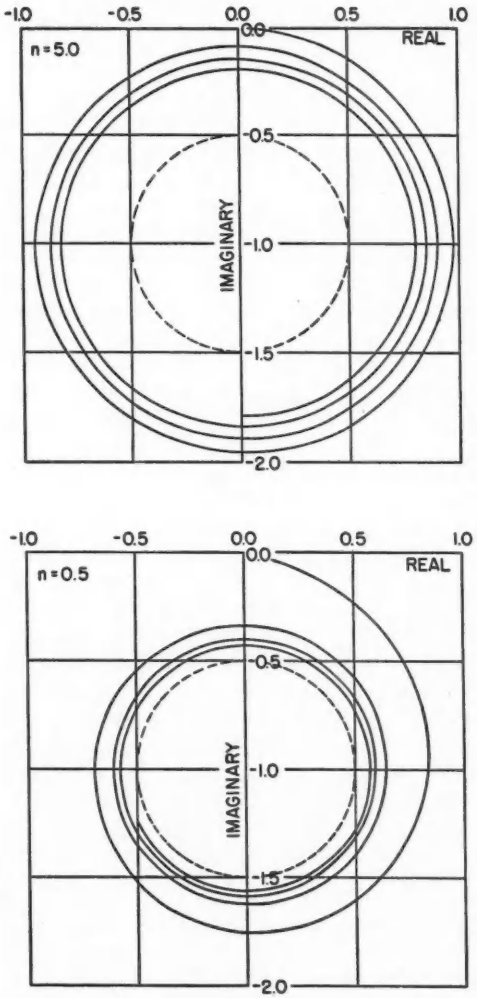


FIG. 2. Plots of the complex amplitudes for values of β from 0 to 7π taken at distances along the axis from the center of the aperture of 5.0 and 0.5 wavelengths.

rapidly upon the circle of radius one-half. As n approaches zero and θ approaches 90 degrees, $u/u_0=\frac{3}{2}$ when an odd number of zones is uncovered and $\frac{1}{2}$ when an even number is uncovered. The best agreement between measurements of microwave diffraction and Eq. (8) is for n close to zero.[1]

Finally, a simple alternative to the patching of Fresnel zone theory in the previous section may be noted. Fresnel's sole purpose in summing the effects from an infinite number of zones was to prove that the amplitude due to the first zone is twice that of the unperturbed wave for points on the axis at a large number of wavelengths from the aperture. The solution of Kirchhoff's integral, Eq. (8) and Fig. 2, yields the same result without reference to the infinitely distant zones, whose individual effects have never been measured. Indeed, Kirchhoff's equation gives all amplitudes in terms of the amplitude of the unperturbed wave. Thus, we may avoid the smoking room argument of whether Kirchhoff or Fresnel was right about the effect of infinitely distant zones. No one has ever measured the intensity at a point relative to which the aperture uncovered more than forty precisely cut zones.

The outstanding fact is that Kirchhoff's equation yields the experimentally measured intensity on the axis of a circular aperture for all points up to and including the center of the aperture itself.

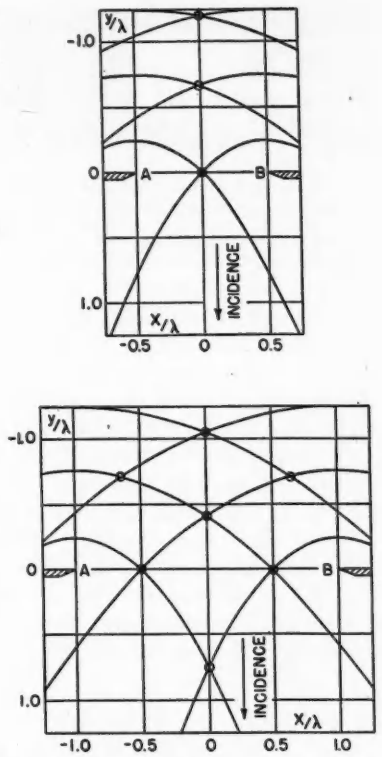


FIG. 3. Families of parabolas whose intersections indicate points of maximum intensity in the diffraction pattern of a slit shown for the cases when the slit is one and two wavelengths wide.

A FORWARD LOOK

Although this paper was written to suggest the patching up of the teaching of century-old treatment of diffraction, some teachers are not averse to pointing out to intermediate students that the theory of diffraction may have to undergo a complete revision. Most outstanding of these disturbers of complacency is R. W. Wood,[8] who notes that the mathematical investigations of Sommerfeld show that Thomas Young's explanation of diffraction is correct. Thomas Young described the diffraction pattern of an aperture as an interference pattern between the geometrically propagated wave through the aperture and waves originating at the edge of the aperture. Thus, the diffraction pattern of a slit upon which a plane wave is incident normally is a simple interference pattern between three waves, a plane wave through the slit and two cylindrical waves from the edges.

Wood[8] shows graphically that the interference pattern between the plane wave and the cylindrical wave from the edge is a family of parabolas with their common principal focus at the diffracting edge. The interference pattern of three waves in a plane perpendicular to the edges is not a set of curves but an array of points where the two families of parabolas intersect as shown in Fig. 3, the points where all three waves are in phase.

Figure 4 is a photograph of a lantern slide holder designed for the projection of two superimposed lantern slides that can be moved over

[8] R. W. Wood, *Physical Optics* (The Macmillan Company, New York, 1933), third edition, pp. 218–221 and Fig. 139.

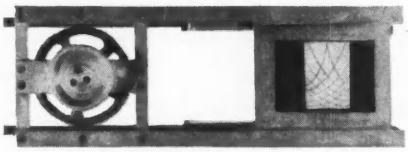


FIG. 4. Lantern slide holder for moving two superimposed lantern slides over each other. As the slit is opened, the points of maximum intensity move from the standing wave region through the slit into the Fresnel region. (Courtesy of The General Electric Company.)

each other. Each slide contains one edge of the slit and its accompanying family of parabolas. As the slit is opened, the points of intersection of the parabolas move from the standing wave region through the aperture into the Fresnel region. Around points where the parabolas intersect at small angles, the contours of intensity are elongated. Where the parabolas intersect at nearly right angles, the contours are nearly circles and the intensity gradients greatest. In the photograph the slit is nearly four wavelengths wide and four peaks of intensity are approaching the plane of the aperture as the slit is being opened. This diagram is the Young counterpart of the Fresnel zone method of predicting the points of maximum and minimum intensity.

In the Fresnel diffraction region the positions of maximum and minimum intensity, as predicted by both Young and Fresnel, are seen to be identical from the geometry; but, in the plane of the aperture where Fresnel assumed constant intensity and phase, Young's theory yields the experimentally observed positions of maxima and minima.

Indeed, the empirical equation[1] which fits the experimentally observed distribution of intensities in a circular aperture is observed to contain the ideas of Thomas Young.

The gaps in experimental knowledge of even the simplest cases of diffraction are numerous and the gap of needed calculations from Sommerfeld's exact electromagnetic theory of diffraction to yield graphical plots of intensity is wide. The problems and the tools are ready for a large number of challenging studies in diffraction for seniors and graduate students.

The study of diffraction is neither cut nor dried.

On the Use of Curve Differentials in Thermodynamics

F. H. CRAWFORD
Williams College, Williamstown, Massachusetts
(Received June 8, 1950)

The question of an appropriate and precise statement of the first law of thermodynamics suitable for infinitesimal reversible changes is approached through a suitable integral form. This latter expresses the work done on and the heat added to the system in terms of "curve" (line) integrals along a specified curve of change, C. As a result, heat functions Q_c and work functions W_c are defined in such a way as to be, on specifically restricted types of curves, single-valued continuous functions of a *single* variable. The first law is then written as $dU = dQ_c + dW_c$, where dU is expressible in terms of *two* independent increments, while the "curve" differentials, dQ_c and dW_c, need but *one*. Applications are made to the heat capacities C_p, C_v, etc., the calculation of second derivatives, etc. In systems with more than two independent variables the concept of "curve" differentials leads at once to "surface" differentials with immediate applications.

IN passing from the statement of the first law of thermodynamics suitable to finite changes to a form appropriate for differential changes, certain difficulties arise. Thus, if U be the internal energy of the system and both heat and work be reckoned positive when they *increase* U, we have for any change from state A to B,

$$U_B - U_A = Q + W. \tag{1}$$

In about two-thirds of the thermodynamics textbooks with which the author is familiar the statement in Eq. (1) is then followed almost immediately by the "differential" form

$$dU = dQ + dW. \tag{2}$$

It is usual to explain, of course, that dU on the one hand and dQ and dW on the other are

different kinds of differentials, the first being a *total* or *exact* differential and the latter two, *incomplete* or *inexact* differentials. Having given this warning, the writers then often proceed glibly to introduce quantities such as $(\partial Q/\partial T)_p$, $(\partial Q/\partial T)_V$, etc., and the properties of these are discussed. These are called the heat capacities C_p and C_v, etc. The student is warned, however, that $\partial^2 Q/\partial T\partial p \neq \partial^2 Q/\partial p\partial T$ and $\partial^2 W/\partial T\partial p \neq \partial^2 W/\partial p\partial T$, since dQ and dW are not perfect differentials. It need hardly be emphasized that such a cavalier treatment of fundamentally troublesome points is hardly likely to inspire confidence or induce clear ideas in readers taking up the subject for the first time.

It is apparent that the trouble lies largely in the notation—the use of the symbol d to indicate quite different operations on the two sides of the equation. To avoid this difficulty special differential symbols have been introduced. Thus we have DQ and DW, dQ and dW, and δQ and δW, etc. This, of course, represents an improvement since the reader is at least warned that he is *not* dealing with ordinary total differentials. The notation, however, does not tell him exactly what he *does* have. It is in this respect that the notation used by Osgood[1] is distinctly superior. Thus, he writes D_sQ and D_sW where these are simply abbreviations for the expressions $(dQ/ds)_c ds$ and $(dW/ds)_c ds$, respectively, where ds is an element of arc length along the curve C representing the path of the reversible change under consideration. If x and y are the coordinates of the plane in which the curve of change C is drawn, then the heat and work elements vanish, not simply because the point B of Eq. (1) has approached infinitely close to point A, but because it has done so *along the curve C*.

We may now, in Osgood's notation, write the first law for any finite reversible change between the points A and B along the curve C as follows:

$$U_B - U_A = \int_{C^A}^{B} D_sQ + \int_{C^A}^{B} D_sW. \qquad (3)$$

If we regard A as some fixed point on the curve C, and B as any variable point, the two integrals on the right of Eq. (3) are functions of the variables locating point B. But since B also lies on the curve C, a *single* variable is sufficient to locate it. Thus, if we agree to measure arc length s from the point A, then as s varies, the integrals each vary in a continuous manner. Each may thus be regarded as a function of the *single* variable s. Likewise, any other suitable variable as x or y may be used where s is readily expressed as $s(x)$ or $s(y)$. (This assumes, of course, that the curve C is not one for which x (or y) is a constant.) If we then choose s as our independent variable we may represent our two integrals, respectively, as $Q_c(s)$ and $W_c(s)$. Here $Q_c(s)$ and $W_c(s)$ are simply the "heat function" and "work function," respectively, for the particular curve C. As such, once C is selected, these are ordinary continuous functions of a *single* variable, and not of *two* independent variables as is the case with the internal energy, where $U = U(x, y)$. They should, therefore, strictly be referred to only as "curve functions," since their form will be determined explicitly by the specific curve envisaged.

We may now rewrite Eq. (3) in terms of these curve functions. If we have the usual system of two independent variables (x, y), this gives

$$U(x, y) = U_0 + Q_c(s) + W_c(s), \qquad (4)$$

where, of course, the curve C must be given either directly as a relation between x and y or in some convenient parametric form.

DIFFERENTIAL CHANGES

We may now differentiate both sides of Eq. (4) and have a form of the first law suitable for differential changes along C. We have

$$dU(x, y) = dQ_c(s) + dW_c(s), \qquad (5)$$

where

$$dU(x, y) = (\partial U/\partial x)_y dx + (\partial U/\partial y)_x dy = dU,$$

and $dQ_c(s)$ and $dW_c(s)$ are ordinary differentials of the curve functions $Q_c(s)$ and $W_c(s)$, and, to emphasize their nature, may be called "curve differentials." They represent exactly the quantities designated by Osgood as D_sQ and D_sW, but the new notation has advantages in use and suggestiveness. This is particularly apparent when C is defined as a horizontal or vertical line, as is usually the case in practice. Thus, if C is defined by $y = \text{constant}$, $s \equiv x$, and we may write $dQ_y(x)$ and $dW_y(x)$.

[1] Osgood, *Advanced Calculus* (The Macmillan Company, New York, 1925), p. 454.

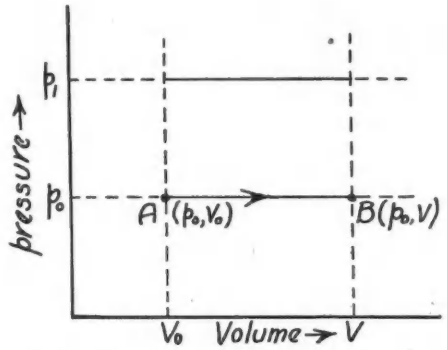


FIG. 1. The heat function, Q_{p0}, is defined as the heat absorbed on moving isobarically from A to B. There is no unique method of connecting Q_{p0} and, say, Q_{p1}, defined for another isobar $p=p_1$. At convenience Q_p may be regarded as a function of the single variable V or T.

In the more general case, C will be defined in such a way that $s=s(y)$, say, where y is a suitable new variable. By *suitable*, we mean that y is (a) not a *constant* for the curve C and (b) a monotone function of s along C. Where the latter condition is not met it is usually satisfactory if we can break our curve into finite segments, on each of which condition (b) holds.

We then have

$$Q_c(s)=Q_c[s(y)]=Q_c'(y),$$

and similarly

$$W_c(s)=W_c[s(y)]=W_c'(y).$$

Here $Q_c(s)$ and $Q_c'(y)$ are heat functions along the same curve and have numerical values identical with those of Q_c and W_c, respectively, at a given point on C. The symbols do, however, represent functions of their respective variables, which are of quite different analytic form. Thus, when s changes from s_0 to s_0+ds and y from y_0 to y_0+dy the changes in $Q_c(s)$ and $Q_c'(y)$ are necessarily equal. In symbols

$$dQ_c(s)=dQ_c'(y),$$

and likewise

$$dW_c(s)=dW_c'(y),$$

whence Eq. (5) becomes

$$dU(x,y)=dQ_c'(y)+dW_c'(y). \tag{6}$$

If then, in accordance with the usual practice in physical discussions, the primes (or whatever designation is employed to indicate change of analytic form with change of variable) are dropped, this gives

$$dU=dQ_c+dW_c \tag{7}$$

as the "curve differential" form of the first law. Here neither the two independent variables implied on the left nor the single independent variable needed on the right is explicitly given. We are free, within the above *implied* limits, to select these variables as convenience dictates.

In practice, we are frequently interested in curves which are specified by holding some variable such as p, V or T constant. Thus we should have heat functions at constant pressure Q_p, at constant volume Q_v, and at constant temperature Q_T. These are, of course, understood to be at some single value of the parameter, such as $p=p_0$, $p=p_1$, etc., giving us Q_{p0}, Q_{p1}, etc. There is no necessary or unique relation between the heat functions along *different* isobars. This results, from the definition of Q_{p0}, as

$$Q_{p0}=\int_{A(p_0,V_0)}^{B(p_0,V)} dQ_p,$$

and there is no way of moving from our fixed reference point A (see Fig. 1) to any other *isobar* while we keep $p=p_0$.

With this in mind we may, for example, take Q_p as a function of T, whence $Q_p(T+\Delta T)-Q_p(T)$ is the heat absorbed along the particular isobar as the temperature changes by ΔT. We then naturally define C_p, the heat capacity at constant pressure, as

$$C_p=\operatorname*{Lim}_{\Delta T\to 0}\frac{Q_p(T+\Delta T)-Q_p(T)}{\Delta T}=\operatorname*{Lim}_{\Delta T\to 0}\frac{\Delta Q_p(T)}{\Delta T},$$

or

$$C_p=dQ_p(T)/dT=dQ_p/dT,$$

where the abbreviated form dQ_p/dT is adequate, since by implication Q_p is a function of T.

On this basis the usual table of heat capacities is made up, where each heat capacity is now defined as an *ordinary* derivative of an appropriate heat function, or simply as the ratio of a curve differential to the appropriate differential of the (single) independent variable. We have

$$C_p=dQ_p(T)/dT=dQ_p/dT.$$

Heat capacity at constant pressure

$C_v = dQ_v(T)/dT = dQ_v/dT.$

Heat capacity at constant volume

$L_v = dQ_T(V)/dV = dQ_T/dV.$

Latent heat of volume expansion

$L_p = dQ_T(p)/dp = dQ_T/dp.$

Latent heat of pressure rise.

There is now no question of trying to explain that $(\partial Q/\partial p)_T$ and $(\partial Q/\partial T)_p$ give us, for some strange reason, L_p and C_p, respectively, although the actual function $Q(T, p)$, which this notation postulates, does not exist. The reason is, of course, that two quite different and unrelated heat functions $Q_T(p)$ and $Q_p(T)$ are involved.

Furthermore, in the case of second derivatives, a similar clarification at once results. We no longer need to insist, for example, that $\partial^2 Q/\partial T\partial p \neq \partial^2 Q/\partial p\partial T$, as the following considerations show. In the first place, C_p is defined at a given point on a given isobar. It should therefore vary, in general, along this isobar and could be regarded as a function of T. If we take a neighboring isobar we should expect a new value of C_p and, hence, essentially expect C_p to vary with both T and p. Since C_p is a perfectly definite physical property of the system we should consequently expect its value to be fixed uniquely by (T, p) and be independent of just how the particular values of T and p were reached. Hence, we write $C_p = C_p(T, p)$ and calculate partial derivatives as for ordinary functions of two variables. Similar considerations apply to L_p. We may therefore consider $(\partial C_p/\partial p)_T$ and $(\partial L_p/\partial T)_p$ to have precise and legitimate meanings. But

$$\left(\frac{\partial C_p}{\partial p}\right)_T = \left[\frac{\partial}{\partial p}\frac{dQ_p}{dT}\right]_T \quad \text{and} \quad \left(\frac{\partial L_p}{\partial T}\right)_p = \left[\frac{\partial}{\partial T}\frac{dQ_T}{dp}\right]_p,$$

and while each of these may be regarded as a *second*-order derivative, since one depends on the heat function $Q_p(T)$ and the other on $Q_T(p)$, there is no reason to presuppose any simple connection, much less an equality.

The ultimate and exact relation of the above two second-order derivatives can be found only after the second law has been introduced and the concept of entropy developed. Then, for any reversible change,

$$dQ_c = TdS,$$

and by the usual limiting process we have

$$C_p = \frac{dQ_p}{dT} = T\left(\frac{\partial S}{\partial T}\right)_p,$$

whence

$$\left(\frac{\partial C_p}{\partial p}\right)_T = \left(\frac{\partial}{\partial p}\frac{dQ_p}{dT}\right)_T = T\left(\frac{\partial^2 S}{\partial p\partial T}\right).$$

Likewise

$$L_p = \frac{dQ_T}{dp} = T\left(\frac{\partial S}{\partial p}\right)_T, \quad \text{and} \quad \left(\frac{\partial L_p}{\partial T}\right)_p = \left(\frac{\partial}{\partial T}\frac{dQ_T}{dp}\right)_p = T\left(\frac{\partial^2 S}{\partial T\partial p}\right) + \left(\frac{\partial S}{\partial p}\right)_T.$$

Consequently, the two second derivatives,

$$\left(\frac{\partial}{\partial p}\frac{dQ_p}{dT}\right)_T \quad \text{and} \quad \left(\frac{\partial}{\partial T}\frac{dQ_T}{dp}\right)_p,$$

differ by exactly $(\partial S/\partial p)_T$.

A precisely parallel situation arises in the case of the work functions at constant T and at constant p ($W_v \equiv 0$ by definition, and is of no interest here), since then

$$dW_c = -pdV.$$

Hence,

$$\frac{dW_p(T)}{dT} = -p\left(\frac{\partial V}{\partial T}\right)_p$$

and

$$\left[\frac{\partial}{\partial p}\frac{dW_p}{dT}\right]_T = -p\left(\frac{\partial^2 V}{\partial p\partial T}\right) - \left(\frac{\partial V}{\partial T}\right)_p,$$

$$\frac{dW_T(p)}{dp} = -p\left(\frac{\partial V}{\partial p}\right)_T \left[\frac{\partial}{\partial T}\frac{dW_T}{dp}\right]_p = -p\left(\frac{\partial^2 V}{\partial T\partial p}\right), \quad \text{etc.}$$

Again there is no question of "$\partial^2 W/\partial p\partial T$" being different from "$\partial^2 W/\partial T\partial p$." These "derivatives" have no meaning while their places are taken by derivatives of two totally different work functions, namely, W_p and W_T.

CYCLIC PATHS

In the above discussion it was explicitly assumed that the curve along which the system moved was a nonreentrant, noncyclic curve. It

is interesting now to examine the situation for a cyclic curve, which we shall assume is not multiply re-entrant but consists of a single closed loop of absolute length s_0 (Fig. 2). Here we take the work function for the curve C to be a function of arc length measured from any convenient point, such as A. Then this work function will take on a continuous range of values as we move around the curve and back to A. If we let s' be the *total* arc length measured from A in a counterclockwise sense, we may write this work function as $W_c'(s')$, the corresponding heat function being $Q_c'(s')$. Then Eq. (4) becomes

$$U(x, y) = U_0(x_0, y_0) + Q_c'(s') + W_c'(s'). \quad (8)$$

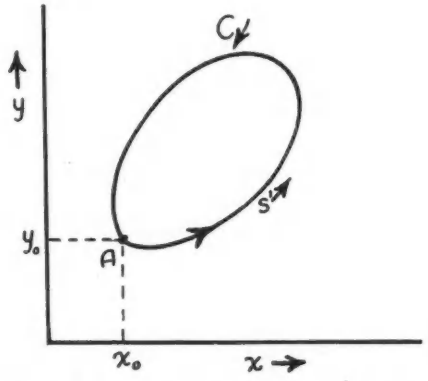


FIG. 2. The closed curve C represents the path of a reversible cycle in the (y, x) plane, where ydx is assumed to represent an element of work. The cycle is assumed of length s_0, while s' is an arc length variable measuring total length of arc from A no matter how many cycles, n, are envisaged.

A disconcerting property of the relations expressed in Eq. (8) is that although U and U_0 by nature are bounded in magnitude, there is no reason in principle why Q_c' and W_c' should be. By executing more and more cycles and thus allowing s' to grow indefinitely, our curve functions must assume greater and greater numerical values. Of course, one will be growing in a positive direction and the other in a negative direction, depending on the sign of s' (and hence the sense of the cyclic motion). And this, quite aside from the question of the physical reality of infinite processes, suggests that a modification is desirable. This modification is readily made as follows. If we agree to subtract from s' the actual length, ns_0, of n completed cycles, we may introduce a new variable s where

$$s = s' - ns_0, \quad 0 \leq s \leq s_0$$

and

$$ds = ds'.$$

Then

$$W_c'(s) = W_c'(s + ns_0) = W_c(s) + nW_{c0}, \quad (9)$$

say, where W_{c0} (the absolute value of the area of the loop) is the constant amount of work *added* to the system per cycle, and n is an integer representing the number of cycles executed. Here n is reckoned positive for counterclockwise and negative for clockwise cycles, making nW_{c0}, therefore, a *constant* for all points on C as close as we choose to go to A. We write likewise

$$Q_c'(s') = Q_c'(s + ns_0) = Q_c(s) - nQ_{c0}, \quad (10)$$

where the minus sign occurs in front of nQ_{c0}, because in counterclockwise motion around the loop, the system *loses* heat Q_{c0} per cycle.[2]

Substitution of Eqs. (9) and (10) in Eq. (8) gives

$$U(x, y) = U_0(x_0, y_0) + Q_c(s) + W_c(s) + n(W_{c0} - Q_{c0}), \quad (11)$$

where (x_0, y_0) are the coordinates of point A. Now let s approach zero from either side of zero and let us agree to *define*

$$Q_c(0) = W_c(0) = 0. \quad (12)$$

In this case Eq. (11) reduces to

$$n(W_{c0} - Q_{c0}) = 0, \quad n \neq 0$$

and hence $W_{c0} = Q_{c0}$. Thus, in general, the last term in Eq. (11) drops out and Eq. (11) reduces to the form appropriate to simple noncyclic paths. Further, we may expect each remaining variable term in Eq. (11) to have a unique derivative at all points of C with the exception of A itself. But this means that Eq. (7) is ap-

[2] It should be pointed out here that the change from $Q'(s')$ to $Q_c(s)$ is simply to avoid the effect of the arbitrary nature of n, the number of cycles executed. Each of these functions is satisfactory on the basis of single valuedness. It is only when we attempt to express them as functions of y (or x) that we obtain multiple-valued functions. In that case, for our simple loop, $Q_c(s(y))$, for example, is a *doubly valued* function of y, while $Q_c(s)$ is a single-valued function of s. Consequently, by drawing tangents to the loop at the two extremes of y we separate it into two simple portions, each with its own monotone relation. The function $Q_c'(s'(y))$, on the other hand, has a multiple valuedness of $2n$, where n is now essentially arbitrary.

plicable to cyclic curves as well as simple nonreentrant curves and holds at all points of these curves including those as close to A as we may choose. The formal extension of the validity of Eq. (7) to other types of curves of change can then be made without serious restrictions.

SYSTEMS WITH MORE THAN TWO INDEPENDENT VARIABLES

In the case of systems with more than two independent variables, the extension is immediate. As an example, consider a thin rod of length L subject to an extensive force F (gravitational effects being neglected) as well as being under a uniform hydrostatic pressure p. Then we should have

$$dU=dQ_c+dW_c, \tag{13}$$

where

$$dW_c=-pdV+FdL. \tag{14}$$

If, to be specific, we choose (T, p, F) as our independent variable set we should have heat functions $Q_{Vp}(F)$, $Q_{TF}(p)$ and $Q_{Fp}(T)$. The differentials of these would then give the corresponding heat capacities. Thus the last would yield $dQ_{Fp}(T)/dT=C_{Fp}$, the analog of the heat capacity at constant pressure in the pure hydrostatic case. The corresponding work function is $W_{Fp}(T)$, and

$$\frac{dW_{Fp}(T)}{dT}=-p\left(\frac{\partial V}{\partial T}\right)_{Fp}+F\left(\frac{\partial L}{\partial T}\right)_{Fp}, \text{ etc.}$$

We also have the possibility of heat and work functions on isothermal *surfaces*—in contrast to the above functions along fixed curves. For we know by the second law that no net external work can be produced by an *isothermal* cycle. Hence, integrating Eq. (13) for such a cycle, we have

$$\oint dU=0=\oint dQ_T+\oint dW_T.$$

But since $\oint dW_T=0$, then $\oint dQ_T$ must vanish. Since these last integrals must vanish for any closed curves of integration as long as T is fixed, we must conclude that dW_T and dQ_T are each themselves perfect or total differentials. Consequently, there exists a work function W_T and a heat function Q_T each of which must be a single-valued function of *two* independent variables, say x and y, as against U which is of course such a function of the set (T, x, y). Accordingly, we may write, T being *constant*,

$$dW_T=-p(T, x, y)dV(T, x, y)+F(T, x, y)dL(T, x, y).$$

If then we choose (x, y) to be, say, (V, L), we may at once equate cross derivatives and have the well-known result

$$-(\partial p/\partial L)_{TV}=(\partial F/\partial V)_{TL}.$$

The same reasoning applies to dQ_T, the total differential of $Q_T(V, L)$. In fact, using Eqs. (13) and (14), we have (where T is again constant throughout)

$$dQ_T=dU+pdV-FdL,$$

or expanding dU and collecting terms,

$$dQ_T=[(\partial U/\partial V)_{TL}+p]dV+[(\partial U/\partial L)_{TV}-F]dL.$$

By equating cross-derivatives we find

$$\left(\frac{\partial^2 U}{\partial L\partial V}\right)_T+\left(\frac{\partial p}{\partial L}\right)_{TV}=\left(\frac{\partial^2 U}{\partial V\partial L}\right)_T-\left(\frac{\partial F}{\partial V}\right)_{TL},$$

a result identical with the above, since the second derivatives of $U(T, V, L)$ cancel. This result can be obtained directly from the fact that dU itself is a perfect differential, but is obtained as above simply to emphasize the consistency of the present point of view and the clarification of method which it introduces.

The writer has been using the heat and work functions defined along specific curves and their corresponding curve differentials for many years in the presentation of thermodynamics. It has been his experience that the method is simple, direct, and convenient in practice. It represents, further, much more than a slight change of notation, though this feature is practically important, since it is an attempt to face explicitly the question of just *what* the symbols in the differential form of the first law really are. As such, it is of great help in teaching the elements of this, the most general and most difficult of undergraduate subjects in physics.

Electric and Magnetic Forces: A Direct Calculation. I*

WILLIAM FULLER BROWN, JR.
Sun Physical Laboratory, Newtown Square, Pennsylvania
(Received October 23, 1950)

In I, the forces acting on an arbitrary portion of an electrically or magnetically polarized body are calculated without introduction of energy arguments, of an atomic model of the polarization, or of assumptions in regard to the physical significance of the macroscopic field vectors within polarized matter. It is pointed out that a magneto-mechanical formula derived by use of ampèrian currents and one derived by use of magnetic poles will, in general, require different "stress" systems in a given physical situation, because a part of the force that is included in the long-range force in one calculation is attributed to stresses in the other. In the present treatment, dipole-dipole interactions are taken as fundamental; poles and ampèrian currents are regarded as purely formal devices; and the two are treated on an equal footing, so that two equivalent sets of formulas are obtained. In II, the general formulas of I are specialized to the case of a fluid. Formulas relating to electrostriction and to the effect of immersing a body in a fluid are derived; their relation to the traditional formulas (based on energy arguments) is discussed.

CONTENTS

Part 1. The forces in general

1.1 Prologue
1.2 Sketch of the procedure
1.3 The forces on a whole body
1.4 The long-range forces on part of a body
1.5 The stresses

Part 2. Fluids and things in them

2.1 Characteristics of a fluid
2.2 Equilibrium relations
2.3 Electrostriction
2.4 Forces on immersed bodies
2.5 The various moments of a magnet

Part 3. Forces on a charged sphere

3.1 Statement of the problem
3.2 Solution of the potential problem
3.3 The long-range force
3.4 The short-range force
3.5 The total force in various cases

Part 4. Energy, thermodynamic functions, and the elastic solid

4.1 Program and precautions
4.2 Work and energy rates
4.3 Thermodynamic functions in general and for a fluid
4.4 The elastic solid
4.5 Epilogue

* II (Parts 3 and 4) will appear in a later issue of this journal.

LIST OF PRINCIPAL SYMBOLS

$_A$ = (subscript) of a Class A fluid.
$\mathbf{B}$ = magnetic flux density.
c = speed of electromagnetic waves.
$\mathbf{D}$ = electric flux density; here defined (for all units) as $\mathbf{E}+\gamma\mathbf{P}$.
D/dt = rate of change in axes moving and rotating with mass element.
d/dt = rate of change in axes moving but not rotating with mass element ($=D/dt+\boldsymbol{\Omega}\times$, for a vector; $=\partial/\partial t+\mathbf{v}\cdot\nabla$, for a vector or scalar).
$\partial/\partial t$ = rate of change in fixed axes.
$\mathbf{E}$ = electric intensity.
de_{zz}, de_{zy}, etc., = differential elements of strain: $de_{zz}=\epsilon_{zz}dt$.
e_{zz}, e_{zy}, etc., = small strains (in elastic solids).
F, F', F, F' = thermodynamic potentials with polarizations and strains as independent variables.
$\mathbf{F}$ = force; $\mathbf{F}_e(\mathbf{F}_m)$ electric (magnetic) force.
$\mathbf{F}_\nu$ = traction across surface with normal $\mathbf{n}$.
$\mathbf{f}$ = acceleration; f_x, f_y, f_z components of acceleration.
$\mathbf{H}$ = magnetic intensity; here defined (for all units) as $\mathbf{B}-\gamma\mathbf{M}$.
$\mathbf{J}_c$, $\mathbf{J}_m$, $\mathbf{J}_t$ = volume density of conduction current, ampèrian current, and total current, respectively.

$\mathbf{L}$ = torque.
l, m, n = direction cosines of normal $\mathbf{n}$.
$\mathbf{m}$ = magnetic moment.
$\mathbf{M}$ = magnetic polarization (magnetic moment per unit volume; magnetization).
$\mathbf{M}$ = specific magnetic polarization (magnetic moment per unit mass).
m = mass; dm, element of mass.
$\mathbf{n}$ = outward normal to surface.
$_n$ = (subscript) normal component.
$\mathbf{p}$ = electric moment.
$\mathbf{P}$ = electric polarization (electric moment per unit volume).
$\mathbf{P}$ = specific electric polarization (electric moment per unit mass).
p_ν = pressure across surface with normal $\mathbf{n}$; p', p_0, $p(\rho)$, pressures of various kinds.
dp = element of magnetic pole strength.
$\mathbf{r}$ = position vector or mutual distance vector ($=r\mathbf{1}$).
r = magnitude of $\mathbf{r}$.
S = surface; dS, element of surface area.
T = kinetic energy.
t = time.
$_t$ = (subscript) tangential component.
$\mathbf{U}_m, \mathbf{U}_m'$ = "magnetic" energies.
V = potential of the mechanical body force per unit mass, or electrostatic potential (as specified).
V = specific volume ($=1/\rho$).
$\mathbf{v}$ = velocity of a mass element.
v_x, v_y, v_z = components of $\mathbf{v}$.
dW = differential element of work.
X, Y, Z = components of mechanical body force per unit mass.
$X_x\ X_y$, etc., = stress components: X_y is the x-component of stress across a surface with normal along y.
X_ν, Y_ν, Z_ν = components of traction $\mathbf{F}_\nu$ across a surface with normal $\mathbf{n}$.
Z, Z', Z, Z' = thermodynamic potentials with field vectors and strains as independent variables.
α = a constant between 0 and 1, otherwise arbitrary.
γ = a constant dependent on the units: $=4\pi$ for gaussian, $=1$ for Lorentz-Heaviside, $=4\pi\times10^{-7}c^2$ for mks-coulomb.
$\epsilon_{xx}, \epsilon_{xy}$, etc., = velocity strains: $\epsilon_{xx}=\partial v_x/\partial x$; $\epsilon_{xy}=\partial v_y/\partial x+\partial v_x/\partial y$.
κ_e = (relative) dielectric constant ($=D/E=1+\gamma\chi_e$).
κ_m = (relative) magnetic permeability ($=B/H=1+\gamma\chi_m$).
$_\nu$ = (subscript) across a surface with normal $\mathbf{n}$.
ρ = mass density.
ρ_m = volume density of magnetic poles.
ρ_c, ρ_p, ρ_t = volume density of conduction charge, polarization charge, and total charge, respectively.
τ = volume; $d\tau$, element of volume.
χ_e = electric susceptibility ($=P/E$).
χ_m = magnetic susceptibility ($=M/H$).
$\mathbf{\Omega}$ = angular velocity of a mass element ($=\frac{1}{2}\nabla\times\mathbf{v}$).
$d\omega$ = differential element of rotation ($=\mathbf{\Omega}dt$).
ω = small rotation (in elastic solids).
$\mathbf{1}$ = unit vector in the direction of $\mathbf{r}$.

1. THE FORCES IN GENERAL

1.1 Prologue

EXPRESSIONS for the forces acting on electrically or magnetically polarized matter, or on charged or magnetized bodies surrounded by such matter, are usually derived by energy arguments. These are all somewhat unsatisfactory. The best of them arrive at the final result by a formal process in which the physical meaning of the various parts of the force is lost sight of. Many of them, furthermore, depend on assumptions whose legitimacy is questionable—such as the assumption that the integrand, in a volume integral representing the total energy, itself represents energy localized in an element of volume. Often the calculation given is valid only for materials in which the polarization is proportional to the field intensity, and there is considerable confusion about the conditions under which the "constant" of proportionality is really constant. Thus, it is usually assumed at first that the dielectric constant is indeed constant as the field builds up; the results are later applied to a fluid whose dielectric constant is a function of density; and it is then shown that the density is a function of field intensity as well as of pressure. It follows that the previously

assumed process of building up the field, at constant dielectric constant, would require an elaborate and variable pressure distribution, and perhaps body forces as well; yet there was no mention of these in the original calculation, and the self-consistency of the whole argument is never demonstrated. Often, too, the argument rests on a rather literal interpretation of the macroscopic field intensity as an "effective" field intensity, or of poles (or ampèrian currents) as things on which forces act physically—interpretations for which neither macroscopic nor atomic theory affords any rigorous justification.

In the following pages, formulas for the forces will be developed by a method that avoids these difficulties. It makes no use of energy arguments, except at a very late stage in connection with the special properties of elastic solids; and it introduces no special assumptions of the sort just mentioned.

The analysis in Parts 1–3 was worked out in 1946–1948 in connection with the work of the Coulomb's Law Committee of the AAPT. It was circulated to the members of the committee; and because some of the formulas appeared incompatible with accepted ones in the literature, criticism was invited. The ensuing discussion led to a clarification of the apparent discrepancies, which in all cases turned out to be matters of interpretation rather than of fact: the calculation leads, as will be shown, to the same final formulas as more orthodox methods.

The possibility of deriving formulas that are seemingly incompatible, but actually equivalent, arises from the fact that there is no unique way of separating the force exerted by one part of a polarized body on an adjacent part into a long-range term, to be calculated by macroscopic electrostatics, and a short-range term, to be calculated by means of "stresses." Thus, the long-range magnetic force can be calculated by use of a volume dipole density, of volume and surface pole densities, or of ampèrian volume and surface current densities; the corresponding "stresses" are all different. As one must not combine volume pole densities with surface current densities, so one must not combine long-range forces computed by one method with "stresses" corresponding to another. It will turn out that for many purposes the various methods are equivalent; however, such equivalence will not be assumed in advance. Instead, a single point of view will be adhered to throughout; and the equivalence of the others, when it holds, will be demonstrated by mathematical transformation of the formulas to alternative forms.

The point of view adopted is the following. Electrostatic and magnetostatic forces will be considered to be exerted directly by matter on matter, at a distance, not by way of stresses in the "ether"; and in a polarized medium, the element by or on which the force is exerted will be taken to be the dipole. The separation of a force into "long-range" and "short-range" terms requires an arbitrary definition when, for instance, the body of matter τ, the force on which is to be calculated, is itself a part of a larger physical body B and is therefore in direct contact with the part τ' of B that is not included in τ. In this case a "force" can be formally calculated by integrating over τ and over τ' the formula for the magnetic force exerted by a dipole of moment $\mathbf{M}'d\tau'$ on a dipole of moment $\mathbf{M}d\tau$, with $\mathbf{M}$ the macroscopic magnetization (magnetic moment per unit volume) and $d\tau$ the element of volume. The long-range force exerted by τ' on τ is hereafter defined as the force formally calculated by this method. The stresses must then include, besides short-range forces of purely mechanical origin, any short-range deviations of the actual magnetic force from the long-range force as just defined; such deviations must be expected because the continuous dipole density is a poor approximation to the actual distribution of polarized atoms or molecules when the elements $d\tau$ and $d\tau'$ are close together.

With this definition of long-range forces and stresses, the entire calculation could, in principle, be carried out without ever introducing magnetic pole densities or ampèrian current densities, and without ever defining the macroscopic field vectors $\mathbf{E}$, $\mathbf{D}$, $\mathbf{H}$, and $\mathbf{B}$ at points of a polarized medium. The formulas would, however, be cumbersome. The pole and current densities will therefore be used freely, but always merely as mathematical devices, to which no physical significance will be attributed. Similarly, the macroscopic field vectors, at points of a polarized medium, will be used, but always merely as auxiliary quantities, defined mathematically; if

they have any physical significance, it must be deduced from the definition. Whatever formal use may be made of pole and current densities, the terms "long-range force" and "stress" will always have the meanings explained in the last paragraph, unless the contrary is explicitly stated.

The use of the dipole, or electric or magnetic moment, as the basic element implies that a force will be considered as exerted on the matter in a volume τ if it is exerted on a charge attached to an atom with center in τ—even though, at the instant under consideration, that charge has been displaced across the surface bounding τ. Atomic concepts, however, will not be introduced explicitly. The results desired can be derived by macroscopic methods; and it is better to reserve atomic theory for problems to which macroscopic theory cannot provide an answer—such as the theoretical calculation of the dielectric constant and the explanation of ferromagnetism.

To PROFESSOR E. C. KEMBLE, Chairman of the Coulomb's Law Committee, I am greatly indebted—first for providing the stimulus that led to this calculation, and second for critical discussions that were very helpful in clarifying the interpretation of the results. Helpful stimulation and criticism were also provided by PROFESSORS C. C. MURDOCK and D. L. WEBSTER, members of the same committee, and by PROFESSOR JOHN A. ELDRIDGE.

1.2 Sketch of the Procedure

The electric and magnetic calculations are for the most part formally identical; and therefore, only one needs to be carried out in detail. Except in a few special illustrations, the notation and terminology used will be those of the magnetic case. The formulas for the electric case can be obtained by replacement of each magnetic quantity by the electric quantity analogous to it, in the formal or "parallel" correlation (for which **E** is the analog of **H**) discussed in Secs. 2.5–2.6 of the Report of the Coulomb's Law Committee of the AAPT[1] (hereafter cited as "the Report").

The first step (Sec. 1.3) is to review the formulas for the magnetic force and torque exerted on a whole magnetized body (the "test body")

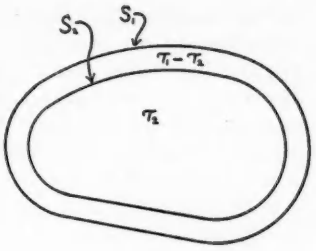


FIG. 1. Method of calculating the force on part of a body.

by sources of field entirely outside it, and separated from it by distances large on the molecular scale. This step is easy.

The second step (Sec. 1.4), which requires more care, is to evaluate the long-range part of the magnetic force and torque exerted on the matter in an arbitrary volume τ of a magnetized body by all sources outside τ, including the rest of the body. In accordance with the definition of the long-range force already outlined, the procedure for calculating it is the following: First, calculate, by use of continuous dipole densities, the magnetic force and torque exerted by everything outside a closed surface S_1 on everything inside a closed surface S_2 completely surrounded by S_1, and separated from it everywhere by a distance molecularly large (Fig. 1); then find the limit of this force and torque as S_1 and S_2 are brought together to form a single surface enclosing a volume τ.

The force and torque calculated in this manner are reliable estimates of physically significant quantities only before the limiting process is carried out; when the distance between S_1 and S_2 becomes of a molecular or atomic order of magnitude, the continuous dipole densities cease to give reliable results. The errors, however, correspond to forces that fall off with distance faster than do dipole forces; because of this short-range character, they can be taken into account, macroscopically, as "stresses." In addition, there are other short-range forces, of nonmagnetic origin, that act across the surface S. In a macroscopic theory there is no possibility of separating these two contributions to the stress—in fact, the separation into two terms will vary with the atomic model adopted. However, regardless of the relative contributions of mechanical and magnetic terms, the resultant stress components must obey certain laws that

[1] Am. J. Phys. **18**, 1–25 and 69–88 (1950).

can be deduced by macroscopic theory alone. The deduction of these laws (Sec. 1.5), culminating in equations of (quasi-static) motion and of equilibrium for an arbitrarily magnetized and deformed body of arbitrary magnetic and mechanical properties, completes the general part of the theory (Part 1).

In Part 2 the special case of fluids will be investigated. It turns out that all the important standard results, including the formulas for electrostriction and for the forces on bodies immersed in a polarizable fluid, can be deduced without use of energy arguments and without introduction of any special assumptions, beyond those implied by the definition of the class of matter being discussed and the assumption that an equilibrium state is possible.

In Part 3 the formulas already derived will be illustrated by working out, in detail, the case of a sphere of arbitrary properties immersed in a Class *A* (as defined in the Report, p. 8) fluid.

In Part 4 energy relations will be touched upon for the first time. The reason for discussing them is that, although they are not necessary for derivation of the general force formulas or of the specialized formulas for a fluid, they are needed for treatment of the interrelations of stress, strain, field intensity, and polarization in a solid substance. This is a large field, and only a few illustrative problems from it will be touched upon.

The equations will be written throughout in a general form, with two constants whose values depend on the units chosen. One of these is the speed of electromagnetic waves, which will be represented by c. The other constant, γ, is determined by the units of force, distance, and charge and is defined by Coulomb's law in the form

$$F_e=(\gamma/4\pi)q_1q_2/r^2. \qquad (1.2\text{–}1)$$

All other quantities are so defined that the equations reduce to those of the gaussian system for $\gamma=4\pi$ and to those of the Lorentz-Heaviside for $\gamma=1$. A system of equations that uses mks mechanical units and the coulomb can be obtained by setting $\gamma/4\pi c^2=10^{-7}$ newton/amp^2. The *gaussian mks system* that results from this procedure uses the practical units for all *circuit* quantities (emf, current, resistance, etc.), but some of the *field* quantities differ from those of the Giorgi system. The equations of this article can be converted to Giorgi-Sommerfeld form by making the substitutions

$$\mathbf{D}=\gamma\mathbf{D}',\ \mathbf{B}=c\mathbf{B}',\ \mathbf{H}=\gamma\mathbf{H}'/c,\ \mathbf{M}=\mathbf{M}'/c \qquad (1.2\text{–}2)$$

and then dropping the primes; the constants γ and c will then occur only in the combinations

$$\gamma=1/\epsilon_0,\quad \gamma/c^2=1/\eta_0=\mu_0. \qquad (1.2\text{–}3)$$

In the present notation

$$\mathbf{D}-\mathbf{E}=\gamma\mathbf{P}, \qquad (1.2\text{–}4)$$

$$\mathbf{B}-\mathbf{H}=\gamma\mathbf{M}; \qquad (1.2\text{–}5)$$

whatever convention may be adopted in regard to the dimensions of charge, all four field vectors ($\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$) have the dimensions of force per unit charge, and both polarization vectors ($\mathbf{P}$, $\mathbf{M}$) have the dimensions of surface charge density.

Apart from this consolidation of units, the notation is that of the Report.

1.3 The Forces on a Whole Body

The magnetic field of the external sources can be described by either of two vectors, $\mathbf{H}_0$ and $\mathbf{B}_0$. The first is calculated by adding to the field intensity of the magnetizing currents, Eq. (2–24) of the Report, the coulomb field intensity of the poles in and on external magnetized bodies, Eq. (3–8). The second is calculated by adding to the field intensity of the magnetizing currents that of the ampèrian currents in and on external magnetized bodies, Eq. (3–9). Since the field of the test body itself is not included in $\mathbf{H}_0$ or $\mathbf{B}_0$, the two are equal throughout the space occupied by it, and the dipole formula (3–7) can also be used.

The magnetic moment of the matter in a volume element $d\tau$ of the test body is $\mathbf{M}d\tau$, where $\mathbf{M}$ is the magnetization in $d\tau$. The force exerted on it by the sources of the external field is $\mathbf{M}\cdot\nabla\mathbf{H}_0\,d\tau$, and there is also a couple $\mathbf{M}\times\mathbf{H}_0\,d\tau$, so that the torque about an arbitrary origin O—with respect to which the position vector of $d\tau$ is $\mathbf{r}$—is $[\mathbf{r}\times(\mathbf{M}\cdot\nabla\mathbf{H}_0)+\mathbf{M}\times\mathbf{H}_0\,]d\tau$. The total force and torque are found by integrating these expressions over the volume τ occupied by the body.

From the point of view adopted here, these formulas are fundamental: The force exerted by

the external sources on the matter in $d\tau$ *is* $\mathbf{M}\cdot\nabla\mathbf{H}_0 d\tau$. This is consistent with the observed behavior of small bodies; but it is, of course, not susceptible of direct demonstration when $d\tau$ is part of a larger body, for there is no way of separating, experimentally, the part of the force exerted by external sources from the part exerted by the rest of the body. The ultimate test of the basic formulas is the experimental adequacy of formulas derived from them.

Alternative expressions for the *total* force and torque on the body can be derived by use of the following vector identities:

$$\int \mathbf{n}\cdot\mathbf{u}\,\mathbf{v}\,dS - \int(\nabla\cdot\mathbf{u})\mathbf{v}\,d\tau = \int \mathbf{u}\cdot\nabla\mathbf{v}\,d\tau, \qquad (1.3\text{–}1)$$

$$-\int(\mathbf{n}\times\mathbf{u})\times\mathbf{v}\,dS + \int(\nabla\times\mathbf{u})\times\mathbf{v}\,d\tau = \int[\mathbf{u}\cdot\nabla\mathbf{v}+\mathbf{u}\times(\nabla\times\mathbf{v})-\mathbf{u}\nabla\cdot\mathbf{v}]d\tau, \qquad (1.3\text{–}1')$$

$$\int \mathbf{n}\cdot\mathbf{u}\,\mathbf{r}\times\mathbf{v}\,dS - \int(\nabla\cdot\mathbf{u})\mathbf{r}\times\mathbf{v}\,d\tau = \int[\mathbf{r}\times(\mathbf{u}\cdot\nabla\mathbf{v})+\mathbf{u}\times\mathbf{v}]d\tau, \qquad (1.3\text{–}2)$$

$$-\int\mathbf{r}\times[(\mathbf{n}\times\mathbf{u})\times\mathbf{v}]dS + \int\mathbf{r}\times[(\nabla\times\mathbf{u})\times\mathbf{v}]d\tau = \int\{\mathbf{r}\times[\mathbf{u}\cdot\nabla\mathbf{v}+\mathbf{u}\times(\nabla\times\mathbf{v}) - \mathbf{u}\nabla\cdot\mathbf{v}]+\mathbf{u}\times\mathbf{v}\}d\tau. \qquad (1.3\text{–}2')$$

Here dS is an element of the surface bounding τ, $\mathbf{n}$ is the outward normal, and $\mathbf{u}$ and $\mathbf{v}$ are any vector functions that satisfy the conditions for validity of the divergence theorem in the region τ. The desired formulas are obtained by setting $\mathbf{u}=\mathbf{M}$, $\mathbf{v}=\mathbf{H}_0=\mathbf{B}_0$. In anticipation of later extensions, the external field intensity will be written $\mathbf{H}_0$ when it is used with the "pole densities" $-\nabla\cdot\mathbf{M}$ and $\mathbf{n}\cdot\mathbf{M}$, and $\mathbf{B}_0$ when it is used with the "ampèrian current densities" $c\nabla\times\mathbf{M}$ and $-c\mathbf{n}\times\mathbf{M}$; and $\nabla\times\mathbf{B}_0$ will be retained in the formulas (although it actually vanishes) as a warning against forgetting $\nabla\times\mathbf{B}$ in later formulas. However, $\nabla\cdot\mathbf{B}_0$ will be omitted, since $\nabla\cdot\mathbf{B}=0$ always and $\nabla\cdot\mathbf{D}=0$ in an uncharged dielectric.

The force $\mathbf{F}$ and torque $\mathbf{L}$ are then given by

$$\mathbf{F} = \int \mathbf{M}\cdot\nabla\mathbf{H}_0 d\tau \qquad (1.3\text{–}3)$$

$$= \int[\mathbf{M}\cdot\nabla\mathbf{B}_0+\mathbf{M}\times(\nabla\times\mathbf{B}_0)]d\tau \qquad (1.3\text{–}3')$$

$$= \int\mathbf{n}\cdot\mathbf{M}\mathbf{H}_0 dS + \int(-\nabla\cdot\mathbf{M})\mathbf{H}_0 d\tau \qquad (1.3\text{–}4)$$

$$= \int(-\mathbf{n}\times\mathbf{M})\times\mathbf{B}_0 dS + \int(\nabla\times\mathbf{M})\times\mathbf{B}_0 d\tau \qquad (1.3\text{–}4')$$

$$\mathbf{L} = \int\mathbf{r}\times(\mathbf{M}\cdot\nabla\mathbf{H}_0)d\tau + \int\mathbf{M}\times\mathbf{H}_0 d\tau \qquad (1.3\text{–}5)$$

$$= \int\mathbf{r}\times[\mathbf{M}\cdot\nabla\mathbf{B}_0+\mathbf{M}\times(\nabla\times\mathbf{B}_0)]d\tau + \int\mathbf{M}\times\mathbf{B}_0 d\tau \qquad (1.3\text{–}5')$$

$$= \int\mathbf{r}\times(\mathbf{n}\cdot\mathbf{M}\mathbf{H}_0)dS + \int\mathbf{r}\times[(-\nabla\cdot\mathbf{M})\mathbf{H}_0]d\tau \qquad (1.3\text{–}6)$$

$$= \int\mathbf{r}\times[(-\mathbf{n}\times\mathbf{M})\times\mathbf{B}_0]dS + \int\mathbf{r}\times[(\nabla\times\mathbf{M})\times\mathbf{B}_0]d\tau. \qquad (1.3\text{–}6')$$

If there are conduction currents of volume density $\mathbf{J}_c$ in τ (but not crossing its surface S), there is an additional term $\int(\mathbf{J}_c/c)\times\mathbf{B}_0 d\tau$ in the force; the corresponding term in the electric case, when there is a conduction charge density ρ_c, is $\int\rho_c\mathbf{E}_0 d\tau$. Insertion of $\mathbf{r}\times$ gives the corresponding terms in the torques. These additional terms will not be needed until Sec. 2.4 and will be omitted for the present.

Formulas (1.3–4) and (1.3–6) would be obtained directly if magnetization were interpreted as a displacement of poles and if the force on $d\tau$ were defined as the force exerted on those poles that happened, at the instant under consideration, to be in $d\tau$. Formulas (1.3–4′) and (1.3–6′) would be obtained directly by a similar calculation based on ampèrian currents. These four

formulas will be treated here as having only mathematical significance, in the sense that they are derived by mathematical transformation from the formulas taken as basic: Eqs. (1.3–3) and (1.3–5), or their equivalents Eqs. (1.3–3′) and (1.3–5′).

1.4 The Long-Range Forces on Part of a Body

In Fig. 1, let $\mathbf{H}_0=\mathbf{B}_0$ be the field intensity of all sources outside S_1, and let the "test body" be the part of the actual body that is inside S_2. Then the force and torque are given by the equations of Sec. 1.3 with the integrals taken over τ_2 and S_2; but now

$$\mathbf{H}_0=\mathbf{H}-\mathbf{H}_{12}, \tag{1.4–1}$$

$$\mathbf{B}_0=\mathbf{B}-\mathbf{B}_{12}, \tag{1.4–1′}$$

where $\mathbf{H}$ and $\mathbf{B}$ are the total intensity and flux density, respectively, and $\mathbf{H}_{12}$ and $\mathbf{B}_{12}$ are the parts of them contributed by the matter in τ_1, calculated at points inside τ_2. Each term in $\mathbf{H}_0$ must be calculated by use of the pole formula, Eq. (3.8) of the Report, and each term in $\mathbf{B}_0$ by use of the ampèrian current formula (3.9); the dipole integral (3.7) is now semiconvergent and can be used only if Kelvin cavities are introduced, a procedure that will be avoided here.

From this point on the entire analysis can be carried out in two equivalent but formally different ways, corresponding to choice of $\mathbf{H}$ or of $\mathbf{B}$ as the field vector to be used. It is important to obtain both sets of results, to avoid any danger of misconstruing one or the other as evidence of the greater physical significance of the corresponding field vector. The derivation will in each case be given in detail only for the $\mathbf{H}$-procedure; the $\mathbf{B}$-procedure is generally analogous, in the "antiparallel" sense of the Report.

When $\mathbf{F}$ and $\mathbf{L}$ have been separated into two terms by use of Eq. (1.4–1), the limiting process $S_1,\ S_2\to S$ may be carried out at once for the term involving the total intensity $\mathbf{H}$. Thus

$$\mathbf{F}=\int \mathbf{M}\cdot\nabla\mathbf{H}d\tau+F_1, \tag{1.4–2}$$

where

$$\mathbf{F}_1=-\lim_{S_1,\,S_2\to S}\left\{\int_{S_2}\mathbf{n}_2\cdot\mathbf{M}_2\mathbf{H}_{12}dS_2 + \int_{\tau_2}(-\nabla_2\cdot\mathbf{M}_2)\mathbf{H}_{12}d\tau_2\right\}. \tag{1.4–3}$$

Now

$$\mathbf{H}_{12}=\frac{\gamma}{4\pi}\left\{\int_{S_1}\mathbf{n}_1\cdot\mathbf{M}_1\frac{\mathbf{1}_{12}}{r_{12}^2}dS_1 + \int_{\tau_1}(-\nabla_1\cdot\mathbf{M}_1)\frac{\mathbf{1}_{12}}{r_{12}^2}d\tau_1\right\}, \tag{1.4–4}$$

where $\mathbf{1}_{12}$ is a unit vector pointing from surface or volume element 1 to surface or volume element 2, and r_{12} is the mutual distance. Insertion of Eq. (1.4–4) in Eq. (1.4–3) shows, since $\mathbf{1}_{12}=-\mathbf{1}_{21}$, that $\mathbf{F}_1$ can also be written

$$\mathbf{F}_1=+\lim_{S_1,\,S_2\to S}\left\{\int_{S_1}\mathbf{n}_1\cdot\mathbf{M}_1\mathbf{H}_{21}dS_1 + \int_{\tau_1}(-\nabla_1\cdot\mathbf{M}_1)\mathbf{H}_{21}d\tau_1\right\}, \tag{1.4–5}$$

(i.e., the coulomb forces between the pole distributions 1 and 2 obey the law of action and reaction). If the two expressions (1.4–3) and (1.4–5) are added and divided by 2, and the limiting process is then carried out, the volume term vanishes in the limit; but the surface term does not, for $\mathbf{H}_{12}$ is always evaluated *inside* S_1, whereas $\mathbf{H}_{21}$ is always evaluated *outside* S_2. The limit is

$$\mathbf{F}_1=\frac{1}{2}\int\mathbf{n}\cdot\mathbf{M}(\mathbf{H}_1{}^+-\mathbf{H}_1{}^-)dS, \tag{1.4–6}$$

where $\mathbf{H}_1$ is the coulomb field intensity of the poles associated with the magnetized matter in τ, and $+$ and $-$ denote the values just outside and just inside S.

The discontinuity in $\mathbf{H}_1$ across the surface S affects the normal component only and is of amount $\gamma\mathbf{n}\cdot\mathbf{M}=\gamma M_n$. Therefore,

$$\mathbf{H}_1{}^+-\mathbf{H}_1{}^-=\mathbf{n}\gamma M_n, \tag{1.4–7}$$

and

$$\mathbf{F}_1=\frac{1}{2}\gamma\int\mathbf{n}M_n{}^2dS. \tag{1.4–8}$$

The total long-range force exerted on the matter in τ by everything outside it is, finally,

$$\mathbf{F}=\int\mathbf{M}\cdot\nabla\mathbf{H}d\tau+\frac{1}{2}\gamma\int\mathbf{n}M_n{}^2dS. \tag{1.4–9}$$

The analogous $\mathbf{B}$-calculation is somewhat more troublesome, but the result can be obtained

directly from Eq. (1.4–9) by setting $\mathbf{H}=\mathbf{B}-\gamma\mathbf{M}$:

$$\mathbf{F}=\int[\mathbf{M}\cdot\nabla\mathbf{B}+\mathbf{M}\times(\nabla\times\mathbf{B})]d\tau - \frac{1}{2}\gamma\int\mathbf{n}M_t^2dS, \quad (1.4\text{–}9')$$

where M_t is the tangential component of $\mathbf{M}$ on S.

A similar calculation leads to torque formulas, which differ from the force formulas (1.4–9) and (1.4–9′) only by the presence of the factor $\mathbf{r}\times$ after each integral sign and of a couple term

$$\int\mathbf{M}\times\mathbf{H}d\tau=\int\mathbf{M}\times\mathbf{B}d\tau.$$

In the final force and torque formulas, the terms containing $\mathbf{H}$ or $\mathbf{B}$ can be transformed by use of Eqs. (1.3–1)–(1.3–2′) to the alternative form involving pole densities or ampèrian current densities.

Because of the peculiar form of the last term in Eq. (1.4–9) or Eq. (1.4–9′), the long-range force cannot be expressed as simply so much per unit volume; it depends also on the shape of the volume considered.

1.5 The Stresses

In the standard analysis of stresses, as given for instance by Love,[2] two assumptions are made that are not justified here. One is that any forces acting on the body, other than surface tractions, are of such nature that the force on a volume element $d\tau$ can be expressed as a force per unit volume multiplied by $d\tau$. This is not true of the long-range force just computed, which contains a shape-dependent term. The other assumption is that the torque per unit volume about a point inside $d\tau$, produced by these forces, vanishes with vanishing $d\tau$. This is not true of the magnetic forces, which produce a couple $\mathbf{M}\times\mathbf{H}d\tau=\mathbf{M}\times\mathbf{B}d\tau$ proportional to $d\tau$. Accordingly, it is not permissible to take over the standard formulas of elasticity. Instead, it is necessary to repeat Love's analysis with the additional force and couple terms included.

When this is done, the following results are obtained:

(a) The usual relation $Y_z=Z_y$ does not hold but is replaced by

$$Y_z-Z_y=(\mathbf{M}\times\mathbf{H})_x=M_yH_z-M_zH_y. \quad (1.5\text{–}1)$$

Here Y_z is the y-component of force acting across unit area normal to the z-axis, whereas Z_y is the z-component of force acting across unit area normal to the y-axis, and the two must be distinguished. The fact that Y_z and Z_y may be unequal in a polarized medium has been pointed out by Brillouin,[3] who cites earlier work by Henriot.

(b) The usual dependence of the traction across an element of area with normal $\mathbf{n}=(l, m, n)$ upon the tractions across coordinate planes, $X_\nu=X_xl+X_ym+X_zn$, etc., is replaced by

$$X_\nu=[X_x+\tfrac{1}{2}\gamma(M_x^2-M_n^2)]l + X_ym+X_zn, \quad \text{etc.} \quad (1.5\text{–}2)$$

(c) The equations of motion are

$$(\partial/\partial x)(X_x+\tfrac{1}{2}\gamma M_x^2)+(\partial X_y/\partial y)+(\partial X_z/\partial z) + \mathbf{M}\cdot\nabla H_x+\rho X=\rho f_x, \quad \text{etc.}, \quad (1.5\text{–}3)$$

where (X, Y, Z) is the mechanical body force per unit mass, (f_x, f_y, f_z) is the acceleration, and ρ is the mass-density.

(d) At a free surface, $X_\nu=Y_\nu=Z_\nu=0$; and more generally, the values of these quantities at a bounding surface of a body are equal to the tractions applied to the surface from outside.

As in the usual case, the angular acceleration equations are consistent with the equations already deduced but add nothing new [reference 2, article 56(b)].

Equation (1.5–1) holds with $\mathbf{B}$ substituted for $\mathbf{H}$, since $\mathbf{M}\times(\mathbf{B}-\mathbf{H})=\mathbf{M}\times(\gamma\mathbf{M})=0$. The $\mathbf{B}$-form of Eqs. (1.5–2) and (1.5–3) is

$$X_\nu=[X_x-\tfrac{1}{2}\gamma(M_y^2+M_z^2-M_t^2)]l + X_ym+X_zn, \quad (1.5\text{–}2')$$

$$(\partial/\partial x)[X_x-\tfrac{1}{2}\gamma(M_y^2+M_z^2)]+(\partial X_y/\partial y)+(\partial X_z/\partial z) + [\mathbf{M}\cdot\nabla\mathbf{B}+\mathbf{M}\times(\nabla\times\mathbf{B})]_x+\rho X=\rho f_x. \quad (1.5\text{–}3')$$

It has already been mentioned that the separation of the forces into long-range forces and stresses is not unique. The stresses defined here must evidently differ from those defined by any

[2] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Cambridge University Press, New York, 1934), fourth edition, Chapter II.

[3] L. Brillouin, *Les Tenseurs en Mécanique et en Elasticité* (Dover Publications, New York, 1946), pp. 11, 216–217.

energy method, since differentiation of a thermodynamic potential with respect to a strain will automatically give $Y_z=Z_y$. If the long-range forces are computed by associating with each volume element a pole strength $(-\nabla\cdot\mathbf{M})d\tau$ and with each surface element a pole strength $\mathbf{n}\cdot\mathbf{M}dS$, a different set of stresses will be obtained, namely,

$$\bar{X}_x=X_x+\tfrac{1}{2}\gamma M_x^2+H_xM_x, \quad (1.5\text{–}4)$$

$$\bar{Y}_z=\bar{Z}_y=Y_z+H_yM_z=Z_y+H_zM_y. \quad (1.5\text{–}5)$$

For these stresses the equation of motion is

$$(\partial\bar{X}_x/\partial x)+(\partial\bar{X}_y/\partial y)+(\partial\bar{X}_z/\partial z)+(-\nabla\cdot\mathbf{M})H_x+\rho X=\rho f_x, \quad (1.5\text{–}6)$$

and the traction $\bar{\mathbf{F}}_\nu=(\bar{X}_\nu, \bar{Y}_\nu, \bar{Z}_\nu)$ across a surface with normal $\mathbf{n}$ satisfies the usual relation,

$$\bar{X}_\nu=\bar{X}_x l+\bar{X}_y m+\bar{X}_z n. \quad (1.5\text{–}7)$$

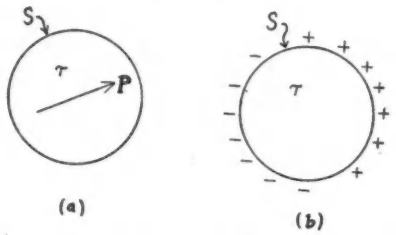


FIG. 2. A spherical portion of a polarized fluid: (a) dipole interpretation; (b) charge interpretation.

The boundary condition at a free surface is

$$\bar{\mathbf{F}}_\nu-(\mathbf{n}\cdot\mathbf{M})\tfrac{1}{2}[\mathbf{H}^+-\mathbf{H}^-]=0. \quad (1.5\text{–}8)$$

If ampèrian currents are used, still a different set of stresses is obtained—namely,

$$\bar{X}_x'=X_x-\tfrac{1}{2}\gamma(M_y^2+M_z^2)+M_yB_y+M_zB_z, \quad (1.5\text{–}4')$$

$$\bar{Y}_z'=\bar{Z}_y'=Y_z-M_yB_z=Z_y-M_zB_y, \quad (1.5\text{–}5')$$

for which the equation of motion is

$$(\partial\bar{X}_x'/\partial x)+(\partial\bar{X}_y'/\partial y)+(\partial\bar{X}_z'/\partial z)+[(\nabla\times\mathbf{M})\times\mathbf{B}]_x+\rho X=\rho f_x. \quad (1.5\text{–}6')$$

The surface traction equation is of the form (1.5–7). At a free surface

$$\bar{\mathbf{F}}_\nu'-(-\mathbf{n}\times\mathbf{M})\times\tfrac{1}{2}[\mathbf{B}^+-\mathbf{B}^-]=0. \quad (1.5\text{–}8')$$

If each of the three sets of stresses is used with the appropriate long-range terms, all three sets of equations are equivalent; but it is clearly not permissible to use one set of stresses with another set of long-range forces.

2. FLUIDS AND THINGS IN THEM

2.1 Characteristics of a Fluid

The object in Part 2 of this study is to apply to the special case of a fluid the general formulas derived in Part 1, and to deduce as many important conclusions as can be deduced by the direct method that is being used, without introduction of energy arguments.

The first problem is to define a fluid. According to Lamb,[4] "The fundamental property of a fluid is that it cannot be in equilibrium in a state of stress such that the mutual action between two adjacent parts is oblique to the common surface." This definition, however, will not suffice in the present problem, for as has been seen the identification of the "stresses" or short-range forces is not unique. If the y-component of traction across a surface normal to z is taken to be the Y_z of Sec. 1.5, then the definition just quoted implies that in a fluid $Y_z=0$, and similarly that $Z_y=0$; then by Eq. (1.5–1) $\mathbf{M}\times\mathbf{H}=0$, so that in a fluid in equilibrium, $\mathbf{M}$ must be in the direction of $\mathbf{H}$; but if the y-component of traction across a surface normal to z is taken to be $\bar{Y}_z$ or $\bar{Y}_z'$, no such conclusion follows. Therefore, Lamb's definition of a fluid, when applied to a polarizable material, becomes ambiguous.

The meaning of this ambiguity, in the analogous electric case, can be understood by considering a spherical region τ within a uniformly polarized fluid (Fig. 2). According to the view adopted here, the electric forces are exerted directly on the dipole moments by other dipole moments and by conduction charges. On this view, the electric field intensity in τ due to distant charges and to polarized matter outside τ is to be computed by use of dipole densities and is $\mathbf{E}_0=\mathbf{E}+\frac{1}{3}\gamma\mathbf{P}$; therefore, the long-range electric torque about the center of the sphere is, by definition, $\tau\mathbf{P}\times\mathbf{E}_0=\tau\mathbf{P}\times\mathbf{E}$. Since the short-range forces exerted across the bounding surface S have, by definition, no tangential component in a fluid, they cannot exert a torque about the center; and therefore, equilibrium requires that

[4] Sir Horace Lamb, *Hydrodynamics* (Dover Publications, New York, 1945), sixth edition, p. 1.

P be along **E**. However, according to the alternative view based on charge densities, there is no long-range electric torque on τ, because the charge density $-\nabla\cdot\mathbf{P}$ vanishes when the polarization is uniform; there is, therefore, on this view, nothing in the definition of a fluid that prevents **E** and **P** from having different directions.

If the spherical surface S were an actual material boundary, both views would lead to the same expression for the torque on the whole body of fluid; it would be considered as exerted on the dipoles according to the first view, Fig. 2(a), and on the surface charges according to the second, Fig. 2(b). When, however, the spherical surface merely identifies a part of the fluid, these surface charges are considered on the second view to have passed out of τ, and electric forces exerted on them are not considered to be exerted on τ. Instead, they are considered to be exerted on charges just outside τ and to be transmitted mechanically by them to matter just inside τ, so that they are included in the stresses.

Because of this non-uniqueness in the identification of "stress," Lamb's definition is not satisfactory for the present purpose. It is clear, however, that there is no uncertainty about interpreting the term "fluid" if S in Fig. 2 is an actual solid boundary enclosing the fluid. Then either view leads to the conclusion that the bounding wall can exert no couple on the fluid, and that therefore **P** must have the direction of **E**.

Accordingly, a fluid will be defined as a material incapable, in equilibrium, of exerting tangential tractions on a solid surface with which it is in contact, or of having such tractions exerted on it. The properties that follow from this definition are independent of the particular interpretation of "stress" adopted.

From this definition, applied to a surface normal to a coordinate axis, it follows at once that $Y_z=Z_y=0$, etc.; and then, by Eq. (1.5–1), that (in the magnetic case) **M** must be along **H**. Then for an arbitrarily oriented surface, Eq. (1.5–2) shows—since (X_ν, Y_ν, Z_ν) must be along (l, m, n)—that

$$X_x+\tfrac{1}{2}\gamma M_x^2 = Y_y+\tfrac{1}{2}\gamma M_y^2 = Z_z+\tfrac{1}{2}\gamma M_z^2 \equiv -p'. \quad (2.1\text{–}1)$$

The vector force per unit area exerted across dS is then

$$\mathbf{F}_\nu = -(p'+\tfrac{1}{2}\gamma M_n^2)\mathbf{n} \equiv -p_\nu\mathbf{n}, \quad (2.1\text{–}2)$$

and the equation of motion (1.5–3) reduces to

$$-\nabla p'+M\nabla H+\rho\mathbf{F}=\rho\mathbf{f}, \quad (2.1\text{–}3)$$

where $\mathbf{F}=(X, Y, Z)$ is the body force per unit mass, and $\mathbf{f}=(f_x, f_y, f_z)$ is the acceleration. The replacement of $\mathbf{M}\cdot\nabla\mathbf{H}$ by $M\nabla H$ is justified by the facts that **M** is in the direction of **H** and that $\nabla\times\mathbf{H}=0$, so that $\mathbf{H}\cdot\nabla\mathbf{H}=\tfrac{1}{2}\nabla(\mathbf{H}^2)=\tfrac{1}{2}\nabla(H^2)=H\nabla H$. The equation of motion (2.1–3) holds if the motion is slow enough to justify application of the equilibrium definition of a fluid, i.e., neglect of ordinary viscosity and also of magneto- or electro-viscosity.

The **B** form of the equations is

$$X_x-\tfrac{1}{2}\gamma(M_y^2+M_z^2) = \cdots \equiv -p''(=-p'-\tfrac{1}{2}\gamma M^2), \quad (2.1\text{–}1')$$

$$\mathbf{F}_\nu = -(p''-\tfrac{1}{2}\gamma M_t^2)\mathbf{n} = -p_\nu\mathbf{n}, \quad (2.1\text{–}2')$$

$$-\nabla p''+M\nabla B+\rho\mathbf{F}=\rho\mathbf{f}. \quad (2.1\text{–}3')$$

2.2 Equilibrium Relations

The relations of Sec. 2.1 follow from the mere definition of a fluid. The fluid will now be assumed to be in equilibrium in a conservative mechanical field of force, so that

$$\mathbf{F}=-\nabla V, \quad \mathbf{f}=0, \quad (2.2\text{–}1)$$

where V is the potential of the mechanical body forces. Then Eq. (2.1–3) becomes the equilibrium equation,

$$-\nabla p'+M\nabla H-\rho\nabla V=0. \quad (2.2\text{–}2)$$

Let a new function Z be defined by

$$Z\equiv -p'-\rho V: \quad (2.2\text{–}3)$$

then from Eq. (2.2–2)

$$\nabla Z=-M\nabla H-V\nabla\rho. \quad (2.2\text{–}4)$$

Operation on Eq. (2.2–4) with $(\nabla H)\times$ and on the result with $(\nabla Z)\cdot$ shows that $(\nabla Z)\cdot(\nabla H)\times(\nabla\rho)$ must vanish identically, or

$$\partial(Z, H, \rho)/\partial(x, y, z)=0, \quad (2.2\text{–}5)$$

so that there must be a functional relation connecting Z, H, and ρ. If this relation is

$$Z=Z(H, \rho), \quad (2.2\text{–}6)$$

then from Eq. (2.2–4), or from the equivalent relation,

$$dZ = -MdH - Vd\rho, \qquad (2.2\text{–}7)$$

(where d represents the change upon going from $\mathbf{r}$ to $\mathbf{r}+d\mathbf{r}$), it follows that

$$M = -\partial Z/\partial H, \quad V = -\partial Z/\partial \rho. \quad (2.2\text{–}8)$$

Then from Eqs. (2.2–3) and (2.1–2), the normal pressure across a surface with normal $\mathbf{n}$ is

$$p_n = p' + \tfrac{1}{2}\gamma M_n^2, \qquad (2.2\text{–}9)$$

with

$$p' = -Z - \rho V = -Z + \rho(\partial Z/\partial \rho). \quad (2.2\text{–}10)$$

These equations assume a slightly simpler form if specific volume $V = 1/\rho$ is used as a variable instead of ρ, and if moment per unit mass $M = M/\rho$ is used instead of M. Then if

$$Z = Z/\rho = -p'V - V, \qquad (2.2\text{–}11)$$

Eq. (2.2–4) becomes

$$\nabla Z = -M\nabla H - p'\nabla V. \qquad (2.2\text{–}12)$$

Hence, Z can be expressed as a function of H and V, and

$$M = -\partial Z/\partial H, \quad p' = -\partial Z/\partial V. \quad (2.2\text{–}13)$$

The **B** equations are the same in form, except that H is replaced by B, p' by p'', and Z by

$$Z' = Z - (p'' - p') = Z - \tfrac{1}{2}\gamma M^2, \quad (2.2\text{–}14)$$

by Eq. (2.1–1′). The equations analogous to Eq. (2.2–8) can be obtained directly from Eq. (2.2–14), since

$$\begin{aligned} dZ' = dZ - \gamma M dM &= -MdH - Vd\rho - \gamma M dM \\ &= -Md(H + \gamma M) - Vd\rho \\ &= -MdB - Vd\rho, \qquad (2.2\text{–}7') \end{aligned}$$

and so, if Z' is expressed as a function of B and ρ,

$$M = -\partial Z'/\partial B, \quad V = -\partial Z'/\partial \rho. \quad (2.2\text{–}8')$$

A function Z' for unit mass may be similarly defined.

It will be recognized that the equations just derived are of similar form to ones usually obtained by thermodynamics, although no thermodynamic reasoning has been invoked in the present derivation. It should also be recognized that the presence of polarization necessitates some care in the treatment of pressure. In ordinary hydrodynamics a fluid, defined as a substance capable of exerting only normal tractions across a bounding surface, is found to exert the *same* normal traction across differently oriented surfaces; but in the presence of polarization this is no longer true. By Eq. (2.2–9), the pressure across a surface normal to the polarization vector exceeds that across a surface tangential to it by $\frac{1}{2}\gamma M^2$. The fluid is isotropic with respect to magnetic and mechanical properties as a whole, but with respect to mechanical properties alone it is anisotropic, with symmetry about the polarization direction but with different properties along and across this direction.

The definition of a fluid as a substance in which there is "equal pressure in all directions" is nevertheless sometimes taken as fundamental. This is the definition given, for instance, in Abraham-Becker[5] (p. 95). Later, however (p. 100), it is stated that this "pressure" has significance only as a tool for calculating changes of density in the electric field and tells nothing about the force exerted on the walls of the vessel. The precise meaning of the Abraham-Becker "pressure" will become clear in the next section; it is by no means easy to extract from the original discussion.

The dependence of the pressure on the orientation of dS and the ambiguity of the pressure concept at an internal point of the fluid have been recognized by Guggenheim;[6] the results obtained here are consistent with his. They will shortly be shown to be consistent also with those of Abraham-Becker as far as experimentally meaningful quantities are concerned. The difference in the definition of "pressure" merely illustrates the non-uniqueness of "stress" discussed in Sec. 1.1.

2.3 Electrostriction

In a fluid subject to an electric or magnetic field, the density depends on the field intensity as well as on the "pressure" (however defined); and, conversely, the polarization at a given field intensity depends on the "pressure." The am-

[5] Max Abraham, *The Classical Theory of Electricity and Magnetism*, revised by Richard Becker, translated by John Dougall, (G. E. Stechert and Company, New York, 1932), Chapter V.

[6] E. A. Guggenheim, Proc. Roy. Soc. (London) **A155**, 49–70 and 70–101 (1936).

biguities inherent in the pressure concept, because of non-uniqueness and anisotropy, make it advisable to choose as independent variable not the pressure, which is essentially a tensor,[7] but the density, which is a scalar. The relations involved in this phenomenon, as derived by energy methods, will be found in Abraham-Becker[5] and in Stratton.[8] They will now be derived without use of energy arguments, for the case usually treated, that of a fluid in which the polarization is proportional to the field intensity as long as the density is constant, but in which the constant of proportionality is a function of the density. Since the electric case is the one usually treated, **H** and **M** will be replaced by **E** and **P** in this section; and since **E** or **H** is practically always used as independent variable in describing linear fluids, the **D** equations will be omitted.

The assumption is that

$$P=\chi_e(\rho)E=(1/\gamma)[\kappa_e(\rho)-1]E, \quad (2.3\text{–}1)$$

where χ_e is the suceptibility and κ_e the (relative) dielectric constant. Since by Eq. (2.2–8)

$$P=-\partial Z(E,\rho)/\partial E, \quad (2.3\text{–}2)$$

it follows by integration that

$$-Z=\tfrac{1}{2}\chi_e(\rho)E^2+\psi(\rho). \quad (2.3\text{–}3)$$

Then

$$V=-\partial Z/\partial\rho=\tfrac{1}{2}\chi_e'(\rho)E^2+\psi'(\rho), \quad (2.3\text{–}4)$$

and

$$p'=-Z-\rho V=\tfrac{1}{2}[\chi_e(\rho)-\rho\chi_e'(\rho)]E^2+\psi(\rho)-\rho\psi'(\rho). \quad (2.3\text{–}5)$$

The pressure at zero field intensity and at density ρ is

$$p(\rho)=\psi(\rho)-\rho\psi'(\rho), \quad (2.3\text{–}6)$$

so that $\psi(\rho)$ can be evaluated from the pressure-density relation at zero field. The density at a given field intensity, at any point in the mechanical field of force, is found by solving Eq. (2.3–4) for ρ; under zero force ($V=\text{const.}=V_0$) this equation gives the following relation between the densities at two points where the field intensities are E_0 and E_1:

$$\tfrac{1}{2}[E_1^2\chi_e'(\rho_1)-E_0^2\chi_e'(\rho_0)]=-[\psi'(\rho_1)-\psi'(\rho_0)]. \quad (2.3\text{–}7)$$

However, differentiation of Eq. (2.3–6) gives

$$p'(\rho)=-\rho\psi''(\rho), \quad (2.3\text{–}8)$$

so that

$$\psi'(\rho_1)-\psi'(\rho_0)=-\int_{\rho_0}^{\rho_1}dp(\rho)/\rho, \quad (2.3\text{–}9)$$

and Eq. (2.3–7) becomes

$$\int_{\rho_0}^{\rho_1}dp(\rho)/\rho=\tfrac{1}{2}[E_1^2\chi_e'(\rho_1)-E_0^2\chi_e'(\rho_0)]=(1/2\gamma)[E_1^2\kappa_e'(\rho_1)-E_0^2\kappa_e'(\rho_0)]. \quad (2.3\text{–}10)$$

This is Eq. (14c) of Abraham-Becker (p. 97) if $p(\rho)$ is identified with their p.

The pressure across a surface with normal **n**, at a point where the field intensity is **E**, differs from the pressure at the same density but at zero field by

$$p_\nu-p(\rho)=[p'-p(\rho)]+\tfrac{1}{2}\gamma P_n^2=\tfrac{1}{2}[\chi_e(\rho)-\rho\chi_e'(\rho)]E^2+\tfrac{1}{2}\gamma\chi_e^2(\rho)E_n^2 \quad (2.3\text{–}11)$$

by Eqs. (2.2–9), (2.3–5), and (2.3–6); or since $\chi_e=(\kappa_e-1)/\gamma$,

$$p_\nu-p(\rho)=(1/2\gamma)\{[\kappa_e(\rho)-1-\rho\kappa_e'(\rho)]E^2+[\kappa_e(\rho)-1]^2E_n^2\}. \quad (2.3\text{–}12)$$

This is Eq. (15a) of Abraham-Becker (p. 101) with the same interpretation of p as before.

Thus the Abraham-Becker "hydrostatic pressure" p is the pressure that would exist at the actual density if the field intensity were zero.

The case of an incompressible fluid is handled more easily by using V rather than ρ as independent variable. Then Eq. (2.2–2) shows that (in the electric case) p' is a function of E and V, and $\partial p'/\partial E=P$, $\partial p'/\partial V=\rho$. For an incompressible fluid, $\rho=\text{const.}=\rho_0$, $p'=\rho_0 V+f(E)$, and $f'(E)=P$, so that P depends only on E and not on the mechanical forces. For a linear incompressible fluid $P=\chi_e E$, with χ_e independent of E and of the mechanical forces (but not necessarily independent of temperature: "equilibrium" implies uniform temperature throughout the container). Then $p'=\rho_0 V+\tfrac{1}{2}\chi_e E^2+\text{const.}$; and when no mechanical body forces are present, $p'=p_0+\tfrac{1}{2}\chi_e E^2$, where p_0 is the pressure at a point where

[7] The thesis that "pressure never has direction" [R. D. Summers, Am. J. Phys. **14**, 311–313 (1946)] fares poorly in a polarized fluid.

[8] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), first edition, pp. 149–151.

$E=0$. Thus, finally,

$$p_r = p_0+\tfrac{1}{2}\chi_e E^2+\tfrac{1}{2}\gamma\chi_e{}^2E_n{}^2$$
$$=p_0+(1/2\gamma)(\kappa_e-1)[E^2+(\kappa_e-1)E_n{}^2]$$
$$=p_0+(1/2\gamma)(\kappa_e-1)[E_t{}^2+\kappa_e E_n{}^2]. \quad (2.3\text{–}13)$$

In this case, or if κ_e is independent of density, the potential problem can be solved first and the pressure distribution then calculated; but in general this is not possible.

2.4 Forces on Immersed Bodies

The forces on charged bodies or magnets immersed in a polarizable fluid are of interest for two reasons. One is the frequent occurrence, in textbooks and in articles about units and dimensions, of the dubious equations, Eqs. (2–1) and (2–15), of the Report without adequate discussion of the limitations on their range of validity. The other is the fact that force or torque measurements are used in the experimental determination of the magnetic properties of specimens and that the specimens are sometimes surrounded by a fluid in order to control the specimen temperature. The object of this section is to give a direct derivation of Eq. (2–11) of the Report,

$$\mathbf{F}_e=(\kappa_e)_A\int_X \mathbf{E}\rho_t d\tau, \quad (2.4\text{–}1)$$

for the force exerted on a charged body X immersed in a Class A fluid. In this equation, the region of integration (τ) is bounded by a surface S that surrounds the body X but is just outside it; S therefore lies completely in the fluid. The charge elements $\rho_t d\tau$ include all conduction and polarization charges—point, line, surface, or volume distributions; the first three are here replaced, for analytical simplicity, by continuously varying densities. The surface S itself is purely geometrical; there is no contribution to the integral from it.

The force $\mathbf{F}_e$ consists of two terms. The first, $\mathbf{F}_{e1}$, is the long-range electric force exerted on the matter in and immediately adjacent to X by everything outside it; this can be found by the method of Sec. 1.4, generalized to allow a conduction charge density in τ. The second term, $\mathbf{F}_{e2}$, is the short-range force exerted across the surface S by the fluid immediately outside it; this can be found by integrating the element of force $-p_r\mathbf{n}dS$ over S.

Force $\mathbf{F}_{e1}$ is given by the analog of Eq. (1.3–4) with $-\nabla\cdot\mathbf{P}$ replaced by the total charge density $\rho_t=\rho_e-\nabla\cdot\mathbf{P}$. As before, S is to be replaced initially by two surfaces S_1, S_2 that later will be brought together; both are outside X. When $\mathbf{E}_0$ is replaced by $\mathbf{E}-\mathbf{E}_{12}$, the result is

$$\mathbf{F}_{e1}=\int_\tau \rho_t\mathbf{E}d\tau+\int_S \mathbf{n}\cdot\mathbf{P}\mathbf{E}dS+\mathbf{F}_1, \quad (2.4\text{–}2)$$

where $\mathbf{F}_1$ is given by the analog of Eq. (1.4–3) except that $-\nabla_2\cdot\mathbf{P}_2$ is replaced by ρ_{t2}. This change does not affect the final formula for $\mathbf{F}_1$, and Eq. (2.4–2) becomes

$$\mathbf{F}_{e1}=\int_\tau \rho_t\mathbf{E}d\tau+\int_S \mathbf{n}\cdot\mathbf{P}\mathbf{E}dS$$
$$+\frac{1}{2}\gamma\int_S \mathbf{n}P_n{}^2dS. \quad (2.4\text{–}3)$$

The Class A fluid, to which Eq. (2.4–1) refers, has a susceptibility $(\chi_e)_A$ that is independent of position in the nonuniform field. If $(\chi_e)_A$ were a function of density, the density would vary with E and therefore with position in accordance with Eq. (2.3–4) (with $V=$const.), and so $(\chi_e)_A$ would vary in position, contrary to hypothesis; therefore $(\chi_e)_A$ must be independent of ρ. Then the right member of Eq. (2.3–10) vanishes, so that ρ is also independent of position, and so too is $p(\rho)$. Equation (2.3–12) then reduces to the form Eq. (2.3–13), whether or not the fluid is compressible. The term p_0 contributes nothing to the force, so that

$$\mathbf{F}_{e2}=-\int_S \mathbf{n}p_r dS=-\tfrac{1}{2}(\chi_e)_A\int_S \mathbf{n}E^2dS$$
$$-\tfrac{1}{2}\gamma(\chi_e)_A{}^2\int_S \mathbf{n}E_n{}^2dS. \quad (2.4\text{–}4)$$

The surface S lies in the fluid; therefore $\mathbf{P}$ in Eq. (2.4–3) is the same as $(\chi_e)_A\mathbf{E}$ in Eq. (2.4–4), and in the sum the last terms cancel. Thus,

$$\mathbf{F}_e=\mathbf{F}_{e1}+\mathbf{F}_{e2}=\int_\tau \rho_t\mathbf{E}d\tau$$
$$+(\chi_e)_A\int_S[\mathbf{n}\cdot\mathbf{E}\mathbf{E}-\tfrac{1}{2}\mathbf{n}E^2]dS. \quad (2.4\text{–}5)$$

If the surface integral is transformed to a volume

integral, the integrand is $(\nabla\cdot\mathbf{E})\mathbf{E}+\mathbf{E}\cdot\nabla\mathbf{E}-\frac{1}{2}\nabla(\mathbf{E}^2)$ $=(\nabla\cdot\mathbf{E})\mathbf{E}=\gamma\rho_t\mathbf{E}$, which combined with the integrand of the other term gives an integrand $[1+\gamma(\chi_e)_A]\rho_t\mathbf{E}=(\kappa_e)_A\rho_t\mathbf{E}$. The result is Eq. (2.4–1). By reversal of the transformation, the volume integral can be changed to a surface integral; the result is

$$\mathbf{F}_e=[(\kappa_e)_A/\ ]\int_S[\mathbf{n}\cdot\mathbf{EE}-\tfrac{1}{2}\mathbf{n}E^2]dS, \quad (2.4\text{–}6)$$

which is Stratton's[8] formula (72) (p. 152).

The magnetic calculation analogous to this in the *parallel* sense (**H**-formulas) is possible only when X contains no conduction currents; then the electric $\rho_t=\rho_c-\nabla\cdot\mathbf{P}$ is replaced by the magnetic $\rho_m=-\nabla\cdot\mathbf{M}$, since the analog of ρ_e vanishes. It must be remembered that the surface is to lie entirely in the fluid; the body X must therefore be a whole magnet or none—it must not be just one end of a magnet (cf. the Report, p. 14).

The magnetic calculation analogous to this in the *antiparallel* sense (**B**-formulas) is possible provided any conduction currents in X constitute complete circuits, so that no current crosses the surface S. The result is

$$\mathbf{F}_m=[1/(\kappa_m)_A]\int_X[(\mathbf{J}_t/c)\times\mathbf{B}]d\tau, \quad (2.4\text{–}1')$$

where $\mathbf{J}_t$ is the total current density, made up, in general, of the conduction current density $\mathbf{J}_c$ and the ampèrian current density $\mathbf{J}_m=c\nabla\times\mathbf{M}$.

Expressions for the magnetic force in surface integral form are the parallel analog of Eq. (2.4–6) and the following equivalent **B**-formulas:

$$\mathbf{F}_m=[1/\gamma(\kappa_m)_A]\int_S[\mathbf{n}\cdot\mathbf{BB}-\tfrac{1}{2}\mathbf{n}B^2]dS$$
$$=[1/\gamma(\kappa_m)_A]\int_S[(\mathbf{n}\times\mathbf{B})\times\mathbf{B}+\tfrac{1}{2}\mathbf{n}B^2]dS. \quad (2.4\text{–}6')$$

Expressions for the torque may be obtained by inserting $\mathbf{r}\times$ in the volume integrals in Eqs. (2.4–1) and (2.4–1′) and in the surface integrals in Eqs. (2.4–6) and (2.4–6′). The proofs parallel those for the forces; the extra factor necessitates some additional integrations by parts in the volume-surface transformations, and these can be simplified by introducing cartesian components at appropriate points in the calculation.

2.5 The Various Moments of a Magnet

When a magnet is immersed in a Class A fluid, there is no *volume* distribution of poles or of ampèrian currents in the fluid. The poles or ampèrian currents at the *surface of contact* of the fluid with the magnet may then be treated, for some purposes, as if they belonged to the magnet rather than to the fluid. There are thus three magnetic moments to be distinguished: that computed from the magnetization of the magnet material alone; that computed by using the pole formula for moment and including the poles contributed by the fluid; and that computed by using the current formula for moment and including the ampèrian currents contributed by the fluid.

The three moments are:

$$\mathbf{m}=\int_\tau\mathbf{M}d\tau, \quad (2.5\text{–}1)$$

$$\mathbf{m}_p=\int_{\tau+S^-+S^+}\mathbf{r}dp, \quad (2.5\text{–}2)$$

$$\mathbf{m}_I=\frac{1}{2}\int_{\tau+S^-+S^+}\mathbf{r}\times d(\mathbf{I}/c). \quad (2.5\text{–}2')$$

Here the integrations are over the volume τ of the magnet and the inside and outside, S^- and S^+, of its surface of contact with the fluid. The element of pole strength dp is equal to $-\nabla\cdot\mathbf{M}d\tau$ in τ, to $\mathbf{n}\cdot\mathbf{M}^-dS$ on S^-, and to $-\mathbf{n}\cdot\mathbf{M}^+dS$ on S^+; $\mathbf{n}$ points outward, and $\mathbf{M}^-$ is magnet magnetization whereas $\mathbf{M}^+$ is fluid magnetization. The element of current $d\mathbf{I}$ is equal to $c\nabla\times\mathbf{M}d\tau$ in τ, to $-c\mathbf{n}\times\mathbf{M}^-dS$ on S^-, and to $+c\mathbf{n}\times\mathbf{M}^+dS$ on S^+. If the integral over S^+ is omitted, the three moments become equal; and they are also equal in free space, where $\mathbf{M}^+$ vanishes.

The moments $\mathbf{m}_p$ and $\mathbf{m}_I$ are convenient for calculating **H** and **B** at points far from the magnet. There the contribution of the magnet and the adjacent poles of the fluid to **H** is $[-\mathbf{m}_p+3\mathbf{m}_p\cdot\mathbf{11}]/r^3$, the field of a dipole of strength $\mathbf{m}_p$; and the contribution of the magnet and the adjacent ampèrian currents of the fluid to **B** is $[-\mathbf{m}_I+3\mathbf{m}_I\cdot\mathbf{11}]/r^3$, the field of a dipole of strength $\mathbf{m}_I$. Since $\mathbf{B}=(\kappa_m)_A\mathbf{H}$ in the fluid, it follows that—regardless of the shape and properties of the magnet—

$$\mathbf{m}_I=(\kappa_m)_A\mathbf{m}_p. \quad (2.5\text{–}3)$$

This relation may be verified by direct calculation.

The moments $\mathbf{m}_p$ and $\mathbf{m}_I$ are also convenient for calculating the force and torque exerted on a magnet in a Class A fluid, in an almost uniform external field $\mathbf{H}'$, $\mathbf{B}'=(\kappa_m)_A\mathbf{H}'$. By means of the formulas of Sec. 2.4 it may be shown (cf. the Report, pp. 14–15) that

$$\mathbf{F}_m=(\kappa_m)_A\mathbf{m}_p\cdot\nabla\mathbf{H}'=(1/(\kappa_m)_A]\mathbf{m}_I\cdot\nabla\mathbf{B}', \quad (2.5\text{–}4)$$

$$\mathbf{L}_m=(\kappa_m)_A\mathbf{m}_p\times\mathbf{H}'=[1/(\kappa_m)_A]\mathbf{m}_I\times\mathbf{B}'. \quad (2.5\text{–}5)$$

The relation of $\mathbf{m}_p$ or $\mathbf{m}_I$ to the moment $\mathbf{m}_0$ in free space can be calculated easily in the case of a uniformly magnetized ellipsoidal magnet. Let the magnetization be along a principal axis of the ellipsoid; the solution for the general case can be found from this by superposition. Magnetization M then produces an internal magnetic intensity $-\gamma DM$ along the same axis, where $4\pi D$ is the demagnetizing factor for that axis as it is usually defined in gaussian or electromagnetic units. For a permanent magnet stabilized in the manner described in Appendix B of the Report,

$$M=M_0'+(\kappa_m'-1)H/\gamma, \quad (2.5\text{–}6)$$

where κ_m' is the reversible permeability and M_0' is the extrapolated magnetization at $H=0$. In free space $H=-\gamma DM$, so that the magnetization is

$$M_0=M_0'/[1+D(\kappa_m'-1)]. \quad (2.5\text{–}7)$$

Now if the scalar potential problem is formulated by means of partial differential equations and boundary conditions, first with and then without a surrounding fluid of permeability κ_m, it will be seen that the equations for the former case can be obtained from those for the latter by replacing M_0' and κ_m' by M_0'/κ_m and κ_m'/κ_m, respectively. Therefore, the scalar potential after immersion in the fluid is the same as for a magnet in free space with magnetization

$$M_p=\frac{M_0'/\kappa_m}{1+D(\kappa_m'/\kappa_m-1)}=\frac{M_0'}{\kappa_m+D(\kappa_m'-\kappa_m)}. \quad (2.5\text{–}8)$$

It follows that

$$\frac{m_p}{m_0}=\frac{M_p}{M_0}=\frac{1+D(\kappa_m'-1)}{\kappa_m+D(\kappa_m'-\kappa_m)}, \quad (2.5\text{–}9)$$

where m_0 is the moment before immersion in the fluid.

Equation (2.5–9) is essentially Eq. (B–9) of the Report. It has been given by Diesselhorst in a paper to be discussed in Sec. 4.5. Explicit formulas for m_p/m_0 (Page's $1/\gamma$) have been obtained by direct calculation by Wilberforce[9] for spheres and by Page[10] and Preston[11] for prolate and oblate spheroids, in the special case $\kappa_m'=1$. A simple case is the elliptic cylinder with semiaxes a and b, magnetized along a: $D=b/(a+b)$, $m_p/m_0=(a+\kappa_m'b)/(\kappa_m a+\kappa_m' b)$. This reduces to the needle value $1/\kappa_m$ for $a\gg b$ and to the disk value 1 for $b\gg a$. For two magnets far apart, if the mutually induced moments are neglected, the effect of the fluid on the force and torque is to introduce a factor of the form Eq. (2.5–9) for each magnet and an additional factor κ_m. (See Eqs. (2.5–4) and (2.5–5).)

The introduction of these various moments is a legitimate item of technique in the formal solution of special classes of problem; but before they can be safely defined and used, it is necessary that the basic concepts of magnetization, field intensity, flux density, pole and current densities, and permeability be already defined and available for use. It is illogical and dangerous to make any of these basic concepts dependent on such a special phenomenon as, for instance, the behavior of needle-magnets in a Class A fluid.

The forms Eqs. (2.5–4)–(2.5–5) of the force and torque equations exhibit explicitly, through the factor $(\kappa_m)_A$ or $1/(\kappa_m)_A$, the special properties required of the fluid. This requirement is concealed, although formal simplicity is gained, if the equations are written in alternative forms (such as $\mathbf{L}_m=\mathbf{m}_p\times\mathbf{B}'=\mathbf{m}_I\times\mathbf{H}'$) in which $(\kappa_m)_A$ does not appear explicitly.[12] Whatever form is preferred, it is clear that for every argument that suggests that "the" moment of a magnet should be defined as $\mathbf{m}_p$, there is an equally valid one that suggests that it should be defined as $\mathbf{m}_I$.

---

[9] L. R. Wilberforce, Proc. Phys. Soc. (London) **45**, 82 (1933).

[10] Leigh Page, Phys. Rev. **44**, 112–115 (1933).

[11] Glenn W. Preston, Am. J. Phys. **3**, 136–9 (1950).

[12] Cf. John A. Eldridge, Am. J. Phys. **15**, 390–397 (1947); **16**, 327–335 (1948).

Newtonian Mechanics and the Equivalence of Gravitational and Inertial Mass

LEONARD T. POCKMAN
San Francisco State College, San Francisco, California
(Received January 24, 1951)

By a historical analysis of the work of Newton and some of his critics and by showing how the equivalence of gravitational and inertial mass can be established in an elementary experimental fashion, it is shown that the presentation of newtonian mechanics can be clarified by emphasizing this equivalence and using it as the starting point.

IN nearly all textbooks on elementary physics Newton's three laws of motion are quoted verbatim. The entire emphasis of the accompanying discussion is placed on the *inertial* aspect of mass. The writer believes that a more consistent and clear development of the subject as well as a more accurate and just account of the essential content of Newton's work is made possible by a different approach. This consists in placing *explicit* emphasis, from the beginning, on the equivalence of the inertial and gravitational aspects of mass. In one form or another Newton's implicit assumption of this equivalence underlies all of his work, including his use of the equal-armed balance in his experimental study of the law of conservation of momentum.[1]

In actual scientific practice we have followed Newton and set up our standard system of masses by means of the equal-armed balance. Yet modern writers on newtonian mechanics do not make explicit the basic connecting link which the equivalence of gravitational and inertial mass forms between dynamical experiments on the one hand and the essentially static experiments of the equal-armed balance on the other. In what follows it will be shown how this connection can be effectively brought out.

The work of Newton is not only related to the work of his great forerunners and contemporaries in science, but has its origin in everyday experience as well. The idea of weight is commonplace. Bodies are classified as light or heavy, according to the muscular effort we must exert to lift them or support them. In his everyday problems of construction, haulage, and transportation, primitive man learned about weight in an elementary and practical way. For all their slave power, the ancient Egyptians could not possibly have built the pyramids had they not already acquired a scientific knowledge of statics. The subject of statics was formalized and extended to floating bodies by the great Greek scientist and engineer, Archimedes. The whole science of statics and its counterpart in construction and civil engineering is based on our knowledge of the all-pervading character of weight—the one gravitational aspect of mass which can be directly perceived.

That all bodies have inertia was an intuitive concept long before the times of Galileo and Newton. The use and popular connotation of our words "heft" or "hefty" in contrast to the connotation of "weight" or "heavy," is a layman's reflection of this intuitive awareness of inertia. Nearly everyone, whether scientifically trained or not, has the habit of "hefting" an object by moving it up and down when estimating its weight or of simultaneously hefting two objects, one in each hand, when comparing their weights. This fact can be interpreted as evidence of the awareness that the inertia of a body is proportional to its weight. In fact, it is a common daily observation that if one body is heavier than another, it is also more difficult to set it in motion even when friction is negligible. A simple example would be the comparison of the difficulty of setting first a man and then a boy in motion on the same swing. This intuitive concept of inertia must go back before recorded history to the time when our primitive ancestors first made use of percussion tools and projectiles. We generally associate Aristotle with erroneous ideas about motion and falling bodies; but in a recent article, Carl Boyer[2] points out that Aristotle in his *De Audibilibus* made the following state-

[1] Isaac Newton, *Mathematical Principles of Natural Philosophy and The System of the World*, revised translation with historical appendix by Florian Cajori (University of California Press, 1934), p. 22.

[2] Carl Boyer, Sci. Monthly **60**, 358 (1945).

ment, "When one is running fast it is difficult to divert the whole body from its impetus in one direction to some other movement." This is a truly remarkable anticipation, in rough qualitative form, of Newton's second law.

Man's sense of inertial mass and momentum became much less vague, more quantitative, with the development of the potter's wheel, the bow and arrow, the catapult, and finally, by Galileo's time, firearms. In this connection, it is of interest to note that according to Dampier,[3] Baliani, captain of archers at Genoa, was the first man to make a distinction between mass and weight; he distinguished *moles* from *pondus*. Baliani was a contemporary of Galileo. Dampier also notes that Leonardo, a century earlier, had said, "Every body has a weight in the direction of its movement."

Galileo is the first man whose work and whose writings impress us as distinctly modern in tone. He first clearly formulated the concept of accelerated motion and by a combination of careful observation, ingenious experiment, and mathematical analysis formulated the law of inertia, "···that any velocity once imparted to a moving body will be rigidly maintained as long as the external causes of acceleration or retardation are removed."[4] Clearly, the idea of inertial mass is implicit in this formulation; and equally clearly, this work lays the basis for the coming advances of Huygens and Newton. But nowhere does Galileo bring out the concept of inertial mass explicitly. Work similar to Galileo's on falling bodies was done earlier by the Dutch scientist and engineer, Simon Stevin. And, according to Dampier, the experiments of Stevin and Galileo were repeated with more accuracy by Baliani who correctly explained the slight differences in time of fall of a heavy and a light body as being due to the resistance of the air. Recognizing the next necessary step, Galileo did some beautifully ingenious experiments in an unsuccessful attempt to unravel the laws of percussion or impact.[5]

Huygens advanced well beyond Galileo.[6] He did impact experiments with hard, almost perfectly elastic spheres and formulated the law of conservation of momentum in terms of weight and velocity. He did not explicitly formulate the concept of mass in connection with impact. His formulations were restricted in application to elastic collisions. The nearness of his approach to an explicit formulation of a concept of mass as distinct from weight can be found in his *Horologium Oscillatorium* of 1673 in which he says: "This tendency of which we have spoken is absolutely similar to that with which heavy bodies suspended by a thread tend to descend. Whence we concluded also that the centrifugal forces of unequal bodies moved around equal circumferences with the same speed are among themselves as the weights or solid quantities." Professor Crew in his *Rise of Modern Physics* has remarked that this is probably the earliest distinction between mass and weight.

Huygens' work on impact was communicated to the Royal Society in 1668. Almost simultaneously Wren and Wallis contributed similar papers on the same subject. Wallis' study was on the impact of perfectly inelastic bodies.[7]

The stage is now set for the entrance of Isaac Newton. What did he add? How did he achieve it? Newton's greatness as a theoretical physicist and mathematician has tended to obscure the fact that he was a superb mechanic and a great experimenter. In the application of his dynamical and gravitational theories, Newton was primarily concerned with astronomical problems. Nevertheless, he not only did his classic work on the vast global tides, the polar flattening, and the precession of the equinoxes, but also showed a keen interest in common everyday forces. Perhaps because of this interest the force concept did not perturb Newton as it did Poincaré, the mathematician, when the latter said, "···but how measure force, or mass? We do not even know what they are."[8] This realism is beautifully

[3] W. C. Dampier, "From Aristotle to Galileo," Chapter 2 of *Background To Modern Science*, edited by Needham and Pagel (MacMillan Company, Ltd., London, 1940), p. 37.

[4] Galileo Galilei *Two New Sciences*, translated by Henry Crew and Alfonso DeSalvio (MacMillan Company, Ltd., London, 1933), p. 215.

[5] Ernst Mach, *The Science of Mechanics*, translated by Thomas McCormack (Open Court Publishing Company, LaSalle, Illinois, 1942), ninth edition, p. 400.

[6] For an excellent account of Huygens' work, see A. E. Bell, *Christian Huygens and the Development of Science in the 17th Century* (Longmans, Green, and Company, New York, 1947).

[7] Reference 6, pp. 64, 112, 115.

[8] H. Poincaré, *The Foundaions of Science*, translated by G. B. Halsted (Science Press, New York, 1946), p. 98.

illustrated in the following passage from the *Principia*[9] in which Newton gives an intuitive example of his Third Law: "If you press a stone with your finger, the finger is also pressed by the stone. If a horse draws a stone tied to a rope, the horse (if I may so say) will be equally drawn back towards the stone; for the distended rope by the same endeavor to relax or unbend itself, will draw the horse as much towards the stone as it does the stone towards the horse, and will obstruct the progress of the one as much as it advances that of the other. If a body impinge upon another, and by its force change the motion [momentum] of the other, that body also (because of the equality of the mutual pressure) will undergo an equal change in its own motion, towards the contrary part."

The writer once asked Professor David L. Webster, "How did Schroedinger get his wave equation?" His immediate reply was, "In the same way Newton got his equations—he guessed it." This answer provides an indispensable clue to the understanding of Isaac Newton's contribution. Certainly Professor Webster's answer implies no blind guess, but rather implies that the great and inspired guesses of Newton and Schroedinger were based upon an intimate and profound understanding of pre-existing theories, relevant experiments, and, in the case of Newton, thoughtful experience with ordinary mechanical and gravitational forces.

Newton's work cannot be judged nor the origins of his ideas fully understood by separately evaluating his laws of motion and his universal law of gravitation. According to a legend recounted by Voltaire, it was in 1665 or 1666 when the young Newton of 23 or 24, away from Cambridge because of the plague, suddenly perceived the connection between the falling of an apple in his Woolsthorpe garden and the acceleration of the moon as it travels in its orbit around the earth. Whether or not the legend is exactly true is not important, for like all good legends it is true in essence. It was during this period that Newton perceived the connection between two such widely divergent phenomena as the acceleration of a freely falling body above the surface of the earth and the motion of the moon in its orbit some quarter of a million miles away. Or in Newton's own words written at a later period, "And the same year [1665 or 1666] I began to think of gravity extending to y^e orb of the Moon,—and thereby compared the force requisite to keep the Moon in her orb with the force of gravity at the surface of the earth, and found them answer pretty nearly."[10] The inaccuracies of that time were associated partly with errors in the value of the radius of the earth, i.e., in the then accepted miles per degree of latitude.[11] Although in 1672 Picard communicated his accurate value of 69$\frac{1}{3}$ miles per degree to the Royal Society, Newton did not publish his own findings until 1686, 18 years after the initial discovery. According to the historical research of the astronomer John Couch Adams,[12] this long delay was occasioned by the fact that it was not until 1685 that Newton succeeded in proving that a spherically symmetrical distribution of mass produces a gravitational field outside the sphere exactly the same as if all of its mass were concentrated at the center.[13]

Newton's work, like all great advances in theory, had brought superficially unrelated phenomena under a common explanation. Moreover, the ideas which came to Newton in 1666 were the start of a great revolution in science. Not only did he guess the inverse square character of the law of gravitation at this time, but he became the first person fully to comprehend the dual aspect of mass in its relation to weight on the one hand and its relation to inertia on the other. Inasmuch as the symmetry of the universal law of gravitation implies that the force of earth on moon is equal and opposite to the force of moon on earth, it follows that Newton may be said to have used his laws of mechanics at this time insofar as the interaction of celestial bodies is concerned. The formal statement of the laws of mechanics and their application to collisions between bodies of everyday or labora-

[9] See reference 1, p. 13.

[10] Florian Cajori, "Newton's twenty year's delay in announcing the law of gravitation," a paper in *Sir Isaac Newton 1727–1927* (Williams and Wilkins, Baltimore, 1928), pp. 159–160.

[11] See reference 1, Appendix, note 40, p. 663.

[12] Florian Cajori, *A History of Mathematics* (MacMillan Company, Ltd., London, 1922), p. 200. See also reference 10, pp. 127–187.

[13] Newton's letter of June, 1686, to E. Halley confirms this as a cause of the delay. See H. M. Taylor, "Sir Isaac Newton," *Encyclopedia Britannica*, eleventh edition, p. 587.

tory experience comes with the writing of the *Principia* in 1686. This will be discussed later.

In 1666 and during his lifetime it was not possible for Newton to make a successful application of the law of action and reaction to the problem of lunar motion. The discrepancy in his verification of the inverse square law which arose from the seventeenth century values of g, earth's radius, lunar distance, etc., was of the same order of magnitude as the discrepancy arising from the neglect of the earth's contrary acceleration toward the moon.

In 1685 and 1686 when Newton wrote the *Principia* and his *System of the World*, he made a formal separation of his laws of motion and the universal law of gravitation. The first book of the *Principia* opens with the following definition of mass: "The quantity of matter is the measure of the same, arising from its density and bulk conjointly." If we were to stop here, we might charge Newton with having founded his mechanics on a logical circle. This Poincaré seems to have done if we judge from his following remark: "What is *mass?* According to Newton, it is the product of volume by density,—it would be better to say that density is the quotient of mass by the volume."[14] But Newton did not stop at this point. Immediately following and directly beneath this introductory (and incomplete when taken out of context) definition we find the following: "$\cdots$it is this quantity that I mean hereafter everywhere under the name 'body' or mass. And the same is known by the weight of each for it is proportional to the weight as I have shown by experiments on pendulums very accurately made which shall be shown hereafter." This statement comes very close to being an explicit statement of the equivalence of inertial mass and gravitational mass, for the word "weight" from the discoverer of the universal law of gravitation may surely be interpreted as implying $Gm'M'/r_e^2$, where m' and M' are the gravitational masses of the pendulum bob and the earth, respectively. Thus, since Newton says that the "quantity of matter" (clearly inertial mass from context) is proportional to the weight, he is saying $m \approx m'$ or by proper choice of units that $m = m'$ where m stands for inertial mass.

As will be shown later, the conclusion that $m \approx m'$ was not a logical deduction from experiment, but constituted a very farsighted guess which has been confirmed directly by the classic experiments of Eötvös[15] and first explained by Einstein's general theory of relativity. But first a word about Newton's pendulum experiments.[16] In essence, they showed with an accuracy of about one part per thousand what Stevin, Galileo, and Baliani had shown less exactly—"that (allowance being made for the resistance of the air) all bodies descend through equal spaces in equal times."[16]

Fortunately, a freely falling body near the surface of the earth has so small a mass compared to that of the earth that no account need be taken of the contrary acceleration of the earth. The fact that all such freely falling bodies have the same acceleration when falling at the same point above the earth's surface is consistent with three discoveries of Newton:

$$a = F/m, \tag{1}$$

$$F = Gm'M'/r^2, \tag{2}$$

$$m = m'. \tag{3}$$

Substitution in Eq. (1) from Eqs. (2) and (3) gives

$$a = GM'/r^2; \tag{4}$$

that is, the acceleration of a freely falling body is independent of its mass m and depends only on M', r, and G, in agreement with experience.

Clearly, neither Eq. (1) nor Eq. (2) nor Eq. (3) is a logical consequence of the common acceleration experienced by all freely falling bodies at a given point above the earth. But the simultaneous truth of all three is a *sufficient* reason for the observed phenomenon as shown by Eq. (4). That they are not *necessary* was specifically shown more than 200 years later by Einstein's more exact formulation in general relativity. Equation (1) is of special interest. Here, in its relation to Eqs. (2) and (3), Eq. (1) is more than a definition; within the framework of Newton's formulation, it is a strictly necessary law.

[14] Reference 8, p. 98. For a similar and earlier criticism of Newton's definition of mass made by Mach, see reference 5, p. 237.

[15] R. Eötvös, Ann. Physik Chem. 59, 354–400 (1896).

[16] See reference 1, p. 568.

In the first book of the *Principia*, following the definitions, Newton states his three laws. Actually, newtonian mechanics is completely formulated by the second law ($F=ma$) in conjunction with the third law (action = reaction). The law of conservation of momentum is then simply a logical consequence of these two laws.

It was for this reason that Newton viewed his experiments on two-bodied collisions as well as those performed before him by Wallis, Wren, and Huygens as a verification of his laws of motion.[17] The masses of the bodies used in these experiments were undoubtedly compared with one another by means of the equal-armed balance, an instrument whose use for commercial purposes dates from ancient time. Thus, the system of masses which Newton employed in the experimental verification of the law of conservation of momentum—a dynamical law—was set up on the basis of an essentially static experiment with the equal-armed balance.

Strictly speaking, the use of the equal-armed balance in comparing masses involves the assumption of the equivalence of gravitational and inertial mass. Thus, at the equator approximately 0.3 percent of the pull of gravity on a body is used to keep it in its circular path around the center of the earth each twenty-four hours. The remainder of the gravitational pull is, of course, borne by the pan of the balance. This 0.3 percent contribution of the inertial mass to the equilibrium of a suspended or supported mass at the equator falls off to zero at the poles like $\cos^2\theta$, where θ is the angle of latitude. Consequently, at all points on the earth, except at the two poles, the valid use of the equal-armed ballance in comparing masses depends upon the exact equivalence of gravitational and inertial mass.

Explicitly, Newton's mechanics can be briefly formulated in modern terms as follows:

(*a*) *Mass.* Our everyday ideas about the masses or relative masses of two bodies stem from our direct muscular perception of their relative weights and/or relative inertias. Such comparisons can be made quantitative by comparing all bodies with a standard mass (e.g., 1 kg) by means of an equal-armed (or unequal-armed) balance. The balance can be purely gravitational (pan balance at either pole), mixed gravitational inertial (the usual balance), or purely inertial (to be described later).

(*b*) *Force.* Define force as the product of mass by acceleration:

$$F=ma. \tag{5}$$

(*c*) *Action and reaction.* In any interaction between two bodies, A and B:

$$F_{AB}=-F_{BA}. \tag{6}$$

As pointed out earlier in connection with Newton's work on freely falling bodies, Eq. (5) or Eq. (1) is more than a definition in that context. Here also, this same equation can be looked upon as more than a mere definition because the law of conservation of momentum is a logical consequence of Eqs. (5) and (6). This law was experimentally verified by Newton and his contemporaries and with much more precision by countless experiments and engineering developments since then. In fact, as is well known, a mathematically equivalent statement of Newton's laws of motion consists of

$$F=ma, \tag{7}$$

$$\text{total momentum before collision} = \text{total momentum after collision}. \tag{8}$$

With this formulation, the law of action and reaction, Eq. (6), becomes a logical consequence of Eqs. (7) and (8).

Historically, it is quite clear why Newton chose the former and not the latter mode of formulation. As pointed out earlier, his universal law of gravitation implies the law of action and reaction and his work on lunar motion antedates by several years his work and that of his contemporaries on the law of conservation of momentum. Consequently, it was natural for Newton to found his mechanics on Eqs. (5) and (6) and to interpret the experiments demonstrating the validity of the law of conservation of momentum as supporting evidence for his laws of motion. This is particularly true because neither his study of the moon's motion nor his derivation of Kepler's laws of motion nor his study of freely falling bodies here on earth provided any evidence, direct or indirect, in his time, for the validity of the law of action and reaction or the law of conservation of momentum.

[17] See reference 1, p. 22.

Newton's formulation of mechanics went uncriticized for nearly two hundred years. But in 1862 Ernst Mach began his extended criticism of Newton's formulation. Mach's views on Newton are to be found in his *Science of Mechanics*,[5] which appeared in seven editions during his lifetime.

Mach objected to the use of the equal-armed balance in setting up a system of masses. According to Mach, "···it is possible to give only *one* rational definition of mass and that a purely dynamical definition."[18] He proposed that the ratio of masses of two bodies should be taken as the inverse ratio of the acceleration produced by each upon the other in a simple binary interaction. Or to put it mathematically, Mach said:

Let

$$m_1/m_2 = -a_2/a_1. \tag{9}$$

He then defined force by Eq. (5). Equation (9) can be rewritten as follows:

$$m_1a_1+m_2a_2=0, \tag{10}$$

or more simply as

$$d(m_1v_1+m_2v_2)/dt=0. \tag{11}$$

From Eq. (11), it is clear that Mach defined mass by assuming the law of conservation of momentum holds in the two-body collisions between the standard mass and masses being determined. For masses determined in this dynamical way, the law of conservation of momentum is true by definition for simple head-on collisions between any mass and the standard mass. But for a simple head-on collision between *any* two of these "dynamical" masses the validity of the law of conservation of momentum must be established as Newton and his contemporaries established it—by experiment. Likewise, the generalizations which include the vector summation of momenta in complex collisions can be established only experimentally, no matter how the masses are set up.

Mach's proposal for setting up a system of dynamical masses has been criticized because of the practical experimental difficulties which his method entails. Mach sought to counter this objection by advocating a dynamical comparison of a pair of masses at opposite ends of a rotating system. Such a method could never hope, in general, to compete in accuracy with the equal-armed balance because the dynamical balancing would depend far too much on the shapes and inhomogeneities of the masses. The balancing of masses with the equal-armed balance, on the other hand, is practically independent of shape or inhomogeneity because the earth is so very large compared to a body of laboratory size.

As noted earlier, Mach's first suggestion for the dynamical comparison of masses was by means of collinear interaction between the two bodies, i.e., a head-on collision between them. In principle, though at great cost, this could have been developed into a precision method. But for Newton, with his correct assumption of the equivalence of gravitational and inertial mass, this was unnecessary; the equal-armed balance with its complete action depending on both the gravitational aspect and the inertial aspect of mass was sufficient. Science has used the equal-armed balance from Newton's time to the present. And today with this equivalence, established in many and varied ways and with a firm theoretical explanation of this equivalence provided by Einstein's general theory of relativity, there is even less need to forsake this ancient instrument.

Mach was the first to show that a system of masses could be set up in a strictly dynamical way; but he rejected the idea of equivalence which was necessary to Newton for the derivation of Kepler's laws, the explanation of the precession of the equinoxes and the behavior of freely falling bodies on the earth. Thus, although Mach achieved a rational and purely dynamical definition of mass, he rejected one of the cornerstones of Newton's mechanics, the equivalence of gravitational and inertial mass.

These questions are not only of historical importance. As Newton's mechanics are taught today, the inertial aspect of mass receives almost all the attention. This emphasis stems from several sources. First and perhaps most important, Newton's own formulation of mechanics in the *Principia* is separated from his work on the *universal law of gravitation*. Then there is Mach's long campaign for the priority of "dynamical mass." This still influences some modern teachers. Sometimes the fact that it would be

[18] Reference 5, p. 342.

just as hard to shake a body on the moon as here on earth but much easier to lift it is incorrectly assumed to show the primacy of inertial mass—this would be true only if the gravitational mass of a body is determined by its weight alone. And finally, notwithstanding the revolution recently wrought by general relativity in celestial mechanics, we live essentially in an age of terrestrial mechanics. Engineers are concerned with the dynamical balancing of high speed engines, and other high speed machines, with rockets and jet propulsion. Today's scientists, for the most part, are concerned with collisions between atoms or nucleons and/or the acceleration and deflection of these in a growing host of devices. In these small-scale phenomena, the gravitational aspect of mass is insignificant.

Nevertheless, prior to atomic physics, the foundations of physics have to be laid and this with the help of the meter stick, clock, and the equal-armed gravitational balance. It is not possible to do this consistently without bringing the equivalence of gravitational and inertial mass sharply into focus at the outset. Furthermore, physics, including elementary physics, embraces celestial mechanics where the gravitational aspect of mass predominates.

The modern development of spectroscopy has made it possible to determine the velocity (relative to earth) of each star in an eclipsing binary. By applying Newton's laws of motion, it is then possible to determine the ratio of the inertial masses of the two stars. But only by the simultaneous application of the laws of motion, the universal law of gravitation, and the assumption of the equivalence of inertial and gravitational mass is it possible to calculate the absolute mass in grams of each member of the binary.

Gravitational masses determined in this way might be termed absolute gravitational masses in order to distinguish them from the relative gravitational masses set up with the equal-armed balance. Nevertheless, through G, they too are expressed in grams. Masses determined by means of the equal-armed balance or by means of inertial comparison as proposed by Mach are of necessity compared with an arbitrarily chosen standard of mass, such as the kilogram. On the other hand, in principle, it is possible to set up a system of absolute gravitational masses without reference to such an arbitrarily chosen standard of mass. An absolute standard of mass could be set up by a definition involving only measurements of length and time. Such a standard could be defined as a homogeneous spherical body which causes a body of negligible mass to fall freely with unit acceleration when at unit distance from the standard body. Then by means of the universal law of gravitation the masses of all other homogeneous spherical bodies (or in principle of any known shape) could be determined as multiples of this defined unit of mass once the free fall acceleration which they produce in small bodies at a measured distance is known.

Such a scheme for determining absolute masses need not be restricted in its application to astronomical bodies. Thus, the Cavendish torsion balance which is normally used to determine G and thence the masses in grams of spherical or nearly spherical bodies (earth, sun, planets, double stars) can be used in principle for determining the absolute masses of bodies of laboratory size. And this could be done as well in a laboratory on the moon as in a laboratory on the earth if only meter sticks and clocks were were available to both laboratories. In order to measure the free fall acceleration of the Cavendish beam toward two identical spheres, the acceleration of the beam at each point in its oscillation must first be measured with clock and meter stick with the two spheres far away. The freely falling acceleration of the beam can then be determined by simple addition or subtraction from the measured acceleration of the beam with the spheres in place at measured distances.

Any attempt to regard inertial mass or gravitational mass as having a more fundamental character must in the end be fruitless and puts us in an untenable position when, as teachers, we proceed to point out to our students that inertial masses are chosen by means of the equal-armed *gravitational* balance. Newton correctly assumed the equivalence of gravitational and inertial mass. Eötvös directly proved the equivalence experimentally with high precision,[15] and Einstein in his general relativity has explained this equivalence. Therefore, the writer feels that modern teachers of elementary physics would do well to demonstrate the equivalence of

gravitational and inertial mass in a simple way when the elements of newtonian mechanics are first introduced.

This equivalence has been demonstrated by the writer with the help of two instruments, the conventional equal-armed gravitational balance and what may be called an equal-armed dynamical balance. The dynamical balance consists of a horizontal beam which is set in rotation about a vertical shaft. The beam is supported by means of a ball race fastened to the vertical shaft near the top of the shaft. After the horizontal beam has been driven to a convenient angular velocity and is freely coasting, the vertical shaft is suddenly drawn straight downward by means of a spring arrangement. With the dynamical balance in proper zero adjustment the vertical shaft pulls down, visibly clear of the beam, then stops rather suddenly just before the beam falls back onto the top of the shaft. The dynamical balance is now comparable to an adjusted equal-armed gravitational balance, and there is no more need at this point to go into the theoretical details of design of the dynamical balance than there is to make a corresponding analysis of the adjusted equal-armed gravitational balance.

In order to demonstrate the principle by which a system of masses is set up by either type of balance, three bodies have been used. One body has twice the mass of either of the other two. The two equal masses are first shown to be equal by use of the equal-armed gravitational balance. In the same way, the third body is shown to balance the first two. Thus, if either of the lighter bodies be regarded as the standard unit of mass, it will have been shown how to obtain twice the unit mass and in principle, how to obtain any rational multiple of this unit mass by means of the equal-armed gravitational balance.

In order that the necessarily small size of the horizontal bar of the dynamical balance (compared to radius of the earth) cause no appreciable error stemming from the shapes or inhomogeneities of the three masses, it has proved convenient to cut them all from the same brass cylinder. Each body has an axial hole that fits around the horizontal beam of the dynamical balance and each is provided with a small set screw to fasten it to the beam. Then by a procedure analogous to that used with the equal-armed gravitational balance the masses of the three bodies are compared with one another by means of the dynamical balance. The fact that the dynamical balance gives the same results as the gravitational balance provides direct proof of the equivalence of gravitational and inertial mass.

The complete theory of the equal-armed dynamical balance (or the unequal-armed dynamical balance) can be entered into fully without talking about acceleration or without calling on Newton's laws or the law of conservation of momentum. All that is needed is an elementary knowledge of vector summation and the definition of momentum as the product of mass (to be made specific later) times velocity. Then the fact that the loaded rotating bar drops back on to the vertical shaft shows that the total linear momentum is zero. The contribution of the two masses (attached to the horizontal bar) to the linear momentum of the system is $m_1v_1+m_2v_2$. For the equal-armed dynamical balance $v_1=-v_2$; and since $m_1v_1+m_2v_2=0$, it follows that $m_1=m_2$. A simple extension of this same argument gives the theory of the unequal-armed dynamical balance that $m_1/m_2=r_2/r_1$, where r_1 and r_2 denote the positions of the masses m_1 and m_2.

The writer has found this demonstration to be both a dramatic and meaningful way to introduce the foundations of newtonian mechanics to his students.

It is a pleasure to acknowledge the invaluable assistance received from Mr. E. L. Ung in the design and construction of the dynamical balance.

Vote on the New Constitution and By-Laws

More members voted in the mail ballot on the new constitution and by-laws than in any recent annual election of officers. As of April 6, 1951, the count stood at 887 for adoption and 9 against. Such vigorous interest is reassuring, and the committee (*Chairman*, P. Kirkpatrick) is to be congratulated on its work.—R. F. PATON, *Secretary*.

An Elementary Laboratory in Modern Physics

JOSEPH W. STRALEY, KARL H. FUSSLER, AND PAUL E. SHEARIN
The University of North Carolina, Chapel Hill, North Carolina
(Received June 7, 1950)

An introductory course in modern physics for freshmen has been taught at the University of North Carolina for the past three years. The laboratory work consists of a two-hour period each week. Eleven simple experiments have been prepared which illustrate the more significant phases of the subject. Most of the apparatus was constructed in our own shop. The subject matter illustrated by these experiments includes a study of ionization potentials, cathode rays, design and use of a spectroscope, the spectra of atoms and molecules, the photoelectric effect, cosmic rays, and radioactivity.

THERE has recently been a trend in the direction of including more and more modern physics in the introductory sophomore course. This has been accomplished in a variety of ways. In most cases it has simply amounted to extending the time devoted to this subject. In a few cases that have come to the attention of the authors the third quarter in a three-quarter sequence in introductory physics has been devoted entirely to the teaching of modern physics. The Department of Physics at the University of North Carolina now offers a one-quarter course in modern physics to students having no previous experience in physics.[1] An undetermined number of other schools have deferred extending their work in this field because of the difficulty involved in providing suitable laboratory experiments. It will be the purpose of this paper to describe the laboratory work used in connection with the introductory course at the University of North Carolina with the thought that our experience may be of interest to others who wish to extend their course to include more of this material.

Most of us have become accustomed to teaching this material to our somewhat specialized classes of majors. If we are to meet the needs of the times, modern physics must be taught to students who have no long range professional interest in physics. Instead we must expect to provide instruction to students of political science, sociology, and journalism. Who can say that many of those who have written newsstories since Hiroshima did not need a bit of basic instruction in the fundamentals of modern physics?

It is the belief of the authors that knowledge of basic principles of modern physics must cease to be the exclusive property of the experts. Scientists' organizations have shown that at least a segment of the scientific family agree with this belief. In spite of their work the fact remains that the gap between the scientist and the nonscientist is still too large.

There has always been a lag between the discovery of new principles and the appearance of lucid statements regarding these principles in introductory textbooks. Whatever reasons exist for this lag, whether related to the conservative nature of professors, the fear of book companies that texts which break with the past may not have a satisfactory sale, or the difficulty of interpreting modern developments in the vocabulary of the general public, it is clear that the lag exists. We know, however, of no field of knowledge whose vital importance to society has been demonstrated that has remained the property of the scholar. For this reason it seems clear that the lag noted here is not one which will be of long duration.

That this lag is in fact being closed at the present time is demonstrated by the fact that there have appeared in recent years at least two introductory physics texts largely devoted to modern physics. In addition to these there are a dozen or more books which attempt to explain atomic or nuclear physics to the layman.

A course in modern physics would fall far short of its goal without the laboratory. Because so much of this material deals with that which

[1] K. H. Fussler and J. W. Straley, Am. J. Phys. 16, 362 (1948).

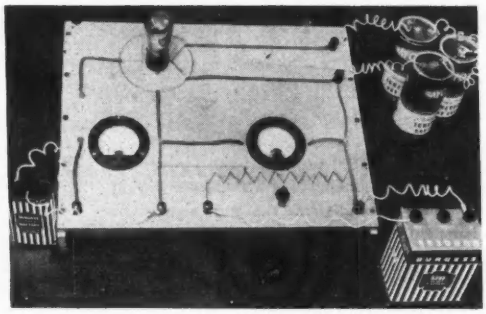


FIG. 1. Apparatus for measurement of ionization potential of a gas.

cannot be seen and because very few of the principles can be related to students' previous experience (as is possible in classical physics), it seems rather more important to have laboratory experiments in modern physics than in the conventional physics courses.

As one might expect, it is impossible to design laboratory experiments in modern physics covering every phase of the subject; but this is a problem in other physics courses as well. However, it has been found possible in the present instance to set up a group of eleven experiments which concern themselves with basic facts of modern physics. In the following sections we will discuss briefly the content of the experiments used in the course at the University of North Carolina and in some cases indicate the apparatus needed to perform these experiments.

STATIC ELECTRICITY

A conventional experiment on static electricity is used as the introduction to the laboratory work. This is justified by the important role which knowledge of static electricity has played in the development of our present concept of the electron. The experiment is also a desirable introductory exercise because it does not require a great deal of classroom instruction in advance of the first laboratory session.

In securing apparatus for this experiment, polystyrene and methacrylate rods have replaced the usual rubber and glass rods because of their superior performance on humid days. In addition silk, wool, pith balls, an electroscope, and a proof plane are provided. In the course of the experiment tests are made which indicate the existence of opposite charges, the difference in the conductivity of various materials, and the technique of charging by induction.

IONIZATION POTENTIAL OF GAS

A photograph of the apparatus used in this experiment is shown in Fig. 1. It will be noted that the essentials of the circuit are indicated on the apparatus being used. The circuit utilizes a type 200 radio tube, which happens to be the simplest gas-filled filament tube we could find. The small retarding potential between plate and grid has the effect of stopping electrons which have experienced an ionizing collision. As the voltage between filament and grid is increased, a somewhat linear relation between this voltage and the plate current is observed up to the ionization potential. At this point there is a

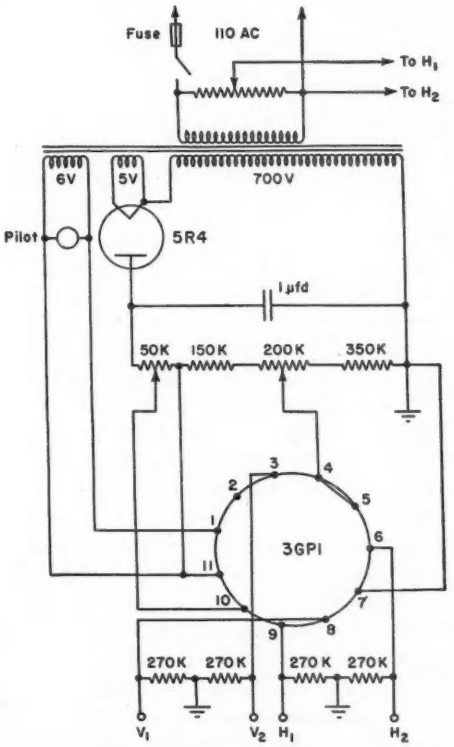


FIG. 2. Circuit of simplified cathode-ray oscilloscope. The bottom view of the oscilloscope tube base is indicated. The pin connections are as follows: 1, 11—heater; 3, 8—vertical deflection plates; 4—focus control; 6, 9—horizontal deflection plates; 7—accelerating electrode; 10—intensity control; 11—cathode.

sharp decrease in the plate current. The tube is then replaced by a type 201, the equivalent vacuum tube, to show that the effect was indeed due to the gas.

Naturally, a number of effects that would be important for a precise measurement of ionization potential are completely ignored in the instructions issued to the students. Contact potentials probably contribute a consistent error to the measurements. Secondary emission could cause a decrease in plate current, but tests have shown that this is not the cause of the decrease observed.

CATHODE-RAY OSCILLOSCOPE

A simple cathode-ray oscilloscope was built in the physics shop utilizing the circuit shown in Fig. 2. The front of the scope was designed as shown in Fig. 3. It will be noticed that the circuit is stripped of all but the bare essentials, that a voltage may be placed on either the parallel or vertical plates at will, and that by merely reversing the switch at the top of the panel one may impress upon the parallel plates a variable ac voltage.

Students performing this experiment start by calibrating the scale using batteries as units of potential, then proceed to associate the sweep produced by a given ac voltage with that read on a voltmeter. From this as a matter of secondary interest they are able to compute the relation between effective and maximum voltage.

PRISM SPECTROSCOPE

The experiment with the prism spectroscope was undertaken primarily to instruct the students regarding the optics utilized in this instrument. A photograph of the apparatus designed for the experiment is shown in Fig. 4.

At the outset of the experiment the student adjusts the collimating lens to produce parallel light. The telescope lens is next adjusted to give a real image on a ground glass screen. The prism is then inserted in the light path. By being mounted on an adjustable table it may be rotated until the ray is caused to pass through at maximum deviation. The image or images may then be examined on the ground glass screen or with a conventional "reading glass" eyepiece.

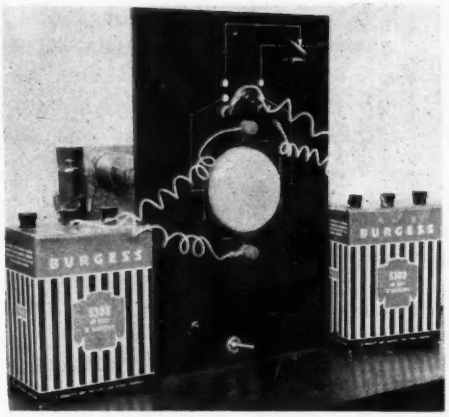


FIG. 3. The laboratory model cathode-ray oscilloscope.

The sources used consist of a germicidal lamp,[2] a fluorescent lamp, and a tungsten lamp, making possible a comparison of the light emitted by each of these sources.

As may be seen in the figure, all the optical parts are mounted on a hinged optical bench. The source may be changed at will from the line source of a straight filament lamp to a slit source of the usual type. Prisms of Lucite, glass, and carbon bisulphide are available for comparison.

DIFFRACTION GRATING

It will be recalled that students taking this course have had little if any previous training in physics. In particular, they have had no previous experience with the wave theory of light. It is the purpose of the experiment with the diffraction grating to supply this lack of previous experience.

The instructions for the experiment refer to the usual analogy of light waves with water

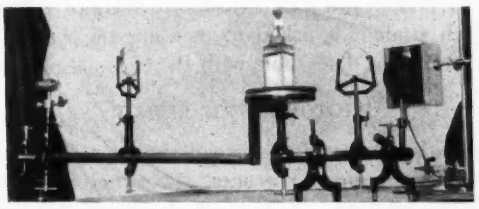


FIG. 4. The model prism spectroscope.

[2] M. S. McCay and E. S. Bishop, Am. J. Phys. 16, 361 (1948).

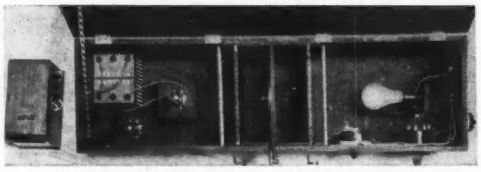


FIG. 5. The photoelectric cell box. The power consumed by the lamp is measured by the wattmeter *W*. Light is rendered parallel by lens L_1 and focused on the photo-cell by lens L_2. The sectored-wheel shutter is located at *S*.

waves and attempt to make realistic the expectation that reinforcements may occur for light diffracted from two small adjacent openings in a transmission grating. A sketch shows the relationship of the direction of reinforcement to the wavelength of light and a table correlates the angle of reinforcement in radians with the ratio of the wave length to the separation of the openings, the grating constant.

The apparatus used consists of a modified version of the conventional diffraction grating apparatus.[3] With this apparatus it is possible for an observer to make a somewhat precise observation of the direction from which diffracted radiation appears to come. This is accomplished by adjusting the position of a needle painted with phosphorescent paint, an adjustment that can be carried out from a position behind the grating. The distance from grating to source is one meter. The needle is carried along the arc of a circle of one-meter radius. Distances along the arc from the slit to the needle measured in meters therefore yield the angle in radians directly.

In addition to gaining experience of the nature of light, students performing this experiment also learn a number of things regarding the nature of a source of light. In particular, they observe the distinction between radiation emitted by tungsten lamps and from various discharge tubes. Each student is asked to measure the extreme wavelengths detectable with the eye.

PHOTOELECTRIC EFFECT

The apparatus for this experiment consists of a 40-watt lamp, a rheostat, a wattmeter, a sectored disk-type shutter, two lenses, and an RCA 929 photoelectric cell installed in a light-tight box. The cell is placed in series with a 90-volt source and a galvanometer (Fig. 5).

Observations are first made to test the linearity of the cell. This is done by noting the deflection of the galvanometer for each of a series of settings of the shutter which, because of its symmetry, can be caused to admit equal increments of light with each of the sequence of adjustments.

Adjustments are then made of the current through the lamp and a correlation is made between the intensity of the light and the power consumed by the lamp.

The Type 929 photoelectric cell is now replaced by the equivalent gas-filled cell, RCA 930. The voltmeter is readjusted to record the same power input as in the first part of the experiment. A measure is made of the current passed by the photoelectric cell when it is irradiated with the same amount of light as in the previous case, i.e., using same shutter adjustment. An approximate "amplification factor" to be associated with the presence of gas in the cell can then be computed.

A STUDY OF SPECTRA

A standard wavelength spectrometer is used in this experiment to make a study of the types of spectra emitted under various circumstances. Several sources are provided, including a tungsten lamp, a sodium lamp, a mercury lamp, and several discharge tubes, some of which contain molecular gases such as I_2 and CO_2. Several bottles containing colored solutions make possible a study of absorption spectra. In particular, a solution of praseodymium nitrate shows several clearcut absorption bands in the visible region of the spectrum.

The student is asked to make a sketch in color of the spectrum which he observes for each of the various sources. Each of the spectra is then to be classified as being of the continuous, line, band, or absorption type.

NATURAL RADIOACTIVITY

In this experiment a gold leaf electroscope with an ionization chamber is utilized to measure the ionizing power of radioactive specimens. A small power supply with variable output voltage enables the student to recharge the electroscope at will by merely pressing a charging key. A

[3] J. Stanley Johnson, Am. J. Phys. 14, 209 (1946).

short focus telescope with an ocular scale and a stopwatch makes possible a measurement of the rate of discharge of the electroscope (Fig. 6).

The first observation is that of noting the normal leakage rate of the electroscope. A number of specimens consisting of carefully weighed amounts of various radioactive ores are then placed in the ionization chamber and the relative activity measured. One specimen of radium is provided. The radium is placed in the ionization chamber, and a sequence of readings is taken in which the specimen is covered by one, two, etc., sheets of onionskin paper. A curve is plotted for rate of fall as a function of numbers of sheets of absorbing paper used. It is found, of course, that a few sheets are sufficient to remove the soft component completely and that any number of sheets will not remove the hard component entirely.

Students, when asked to rate the various experiments from the point of view of general interest, have in most cases picked this one as the most interesting and most instructive experiment in the laboratory.

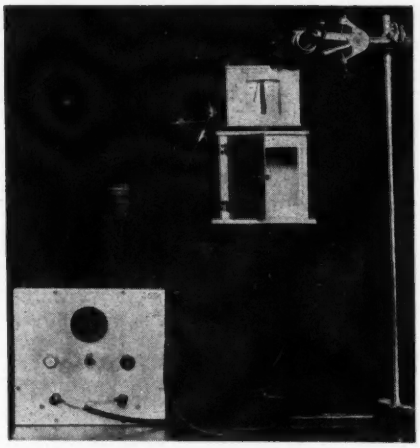


FIG. 6. Ionization chamber for study of natural radioactivity.

GEIGER COUNTERS

Most students find the study of cosmic rays and the use of a Geiger counter to be quite fascinating. In this experiment a simplified Geiger counter is used to study a few characteristics of cosmic rays and the radiation emitted by natural radioactive materials.

When the Geiger counter is first turned on, the random counts due to cosmic activity can be heard and the number of counts per minute is recorded. The tube is then shielded by various thicknesses of lead going up to a thickness of approximately 10 cm. The penetrating power of cosmic rays is quite surprising to students, especially when contrasted with that of the rays emitted by a specimen of uranium oxide, the latter being almost completely absorbed by less than a centimeter of lead.

SPECTRUM OF HYDROGEN

The purpose of this experiment is to familiarize the student with the spectrum of hydrogen and to give him experience in the method used to measure the wavelength emitted by an unknown source. A photograph is provided on which appears the spectrum of hydrogen superimposed upon the spectrum of mercury. This spectrogram was secured by means of a Hilger E1 Spectrograph and, of course, shows reasonably good dispersion. The instructions for the experiment show a drawing of the spectrum of mercury with the "known" wavelengths of the mercury lines. The spectrogram is then placed in a comparator and the position of each of the known and the unknown lines is measured. The wavelength of the lines in the "unknown" hydrogen spectrum is obtained by interpolation. The student is asked to compare the wavelengths secured with those predicted by the Balmer equation.

WILSON CLOUD CHAMBER

It would of course be impossible to provide a cloud chamber with which the tracks of all the various kinds of fundamental particles could be studied. However, a simple cloud chamber showing tracks of alpha-particles may be readily purchased.

In the first part of this experiment the student is given a gallon bottle which contains a few cm^3 of water and to which is attached a simple rubber bulb pump. After the pressure in the bottle has been increased appreciably by means of the pump, the stopper is suddenly removed. Fog is observed to form in the bottle as in a cloud chamber. When smoke or dust is blown into the

bottle, a much denser cloud is observed upon expansion.

A conventional student-model cloud chamber is then used to make observations of the tracks of alpha-particles.

In order to observe the difference between the tracks made by the various kinds of particles a group of photographs taken from the literature is provided which may be studied with a stereoscope. Each of these photographs carries the complete legend telling what sort of track is being observed and what sort of field was imposed on the track, and in other ways helps the student to understand the process taking place. In the write-up the student is expected to answer a number of questions regarding the nature of that which he has seen.

The authors wish to acknowledge the assistance and advice of other members of the Physics Department staff at the University of North Carolina. In particular we wish to express appreciation to Mr. Nestore DiCostanzo of the Physics Shop for his skillful work in the construction of apparatus.

Magnetically Maintained Pendulum

HAROLD P. KNAUSS AND PAUL R. ZILSEL
University of Connecticut, Storrs, Conecticut
(Received November 6, 1950)

An iron bob swinging as a pendulum along the axis of a coil toward and away from its center is maintained in motion by feeding alternating current from a low resistance source into the coil which forms part of a suitably tuned circuit. The iron bob may be replaced by a second coil connected in series. The transfer of energy to the pendulum is shown to be due to the phase difference between the variation in the inductance of the circuit and the corresponding variation in the amplitude of the current with the oscillation of the pendulum: if the variation in the current amplitude lags behind the variation in inductance, the pendulum is driven, otherwise it is damped. In the limit of small amplitude of oscillation the condition for maintenance is that the circuit have predominantly inductive reactance and sufficiently high Q.

A BOB of ferromagnetic material placed near a coil which is part of a suitably tuned circuit is set into vibration along the axis of the coil when an alternating emf is applied to the circuit. J. Bethenod[1] has shown that there exists a solution of the differential equations applying to the problem which yields a force capable of sustaining the oscillation of the pendulum. However, the solution obtained by Bethenod is incomplete, and it clarifies neither the mechanism nor the conditions under which the motion of the pendulum is sustained rather than damped by the magnetic force.*

An experimental study of the pendulum by David Weimer, Alva W. Smith, and the senior author at The Ohio State University in 1941–1942 showed[2] that the maintenance of the motion is due to a time lag in the variation of the current amplitude corresponding to the variation of the self-inductance at any phase of the motion. That is, the current and the attractive force are larger when the bob approaches than when it recedes. It was also shown[3] that nonlinear magnetic properties of iron play no essential part, since the iron bob can be replaced by a second coil suspended coaxially with the first, so that it can swing toward or away from it. The two coils are connected in series so that they attract, and the tuning is adjusted for maximum current when the coils are far apart. Later, it was found that resonance is not essential, but that raw alternating current is sufficient to maintain the motion.

[1] J. Bethenod, Compt. rend. **207**, 847 (1938).

* *Note added in proof:* An application of the magnetically maintained pendulum is described by P. F. Bartunek, Am. J. Phys. **19**, 57 (1951), who remarks on the lack of a theoretical explanation of its operation.

[2] D. Weimer, Phys. Rev. **60**, 65A (1941).

[3] D. Weimer and H. P. Knauss, Phys. Rev. **61**, 544A (1942).

The purpose of the present paper is to consider the conditions which must be satisfied for maintaining the motion, and their effect on the frequency of oscillation of the pendulum. The periodic variation in the self-inductance of the circuit due to the oscillation of the pendulum will be considered as a perturbation on the stationary system, and we shall confine ourselves to a first-order calculation in which the amplitude of oscillation is considered to be small.

The equation for the current is

$$[d(Li)/dt]+Ri+q/C=E(t), \tag{1}$$

where

$$E(t)=E_0\cos(\omega_0 t+\varphi_0) \tag{2}$$

is the applied emf.

The motion of the pendulum satisfies the equation

$$(md^2x/dt^2)+(\beta dx/dt)+k(x+X)=F(t). \tag{3}$$

Here x is the displacement of the pendulum from its equilibrium position, the latter being determined *with the current turned on;* thus, X is the distance between the equilibrium position with current flowing and that without current flowing. The damping constant β of the pendulum applies without current flowing, and

$$F(t)=\tfrac{1}{2}i^2dL/dx \tag{4}$$

is the magnetic force.

We assume that, in first approximation, the solution of Eq. (3) will be

$$x=x_0\sin\omega_1 t, \tag{5}$$

with

$$\omega_1\ll\omega_0. \tag{6}$$

Thus, we are neglecting higher frequency components of x due to the variation of force in Eq. (4) with the frequency ω_0 of the applied emf. Because of the condition Eq. (6), these components will be very small. In general, ω_1 will, of course, differ somewhat from the frequency of the pendulum when there is no current flowing, since, as we shall see, the force $F(t)$ contains a component which is in phase with the displacement of the pendulum.

To carry out the perturbation calculation we rewrite Eq. (1) in the form

$$(L_0di/dt)+Ri+(q/C)-E = -(idL/dt)-(L-L_0)di/dt, \tag{7}$$

where $L_0=L(x=0)$, and the terms on the right of the equation represent the perturbation due to the variation of the self-inductance. We assume a series development for $i(t)$:

$$i=i_0+i_1+i_2+\cdots=(dq_0/dt)+(dq_1/dt)+\cdots, \tag{8}$$

where i_0 is the unperturbed current obtained by setting the right-hand side of Eq. (7) equal to zero, and i_n obeys the equation

$$(L_0di_n/dt)+Ri_n+(q_n/C) = -(i_{n-1}dL/dt)-(L-L_0)di_{n-1}/dt. \tag{9}$$

We shall carry out the solution only to the extent of obtaining i_1.

The solution for the unperturbed current i_0 can be written down immediately:

$$i_0=I_0\cos\omega_0 t;\quad I_0=E_0/Z_0, \tag{10}$$

with

$$Z_0=[R^2-(\omega_0L_0-1/\omega_0C)^2]^{\frac{1}{2}} \tag{11}$$

and the phase angle, φ_0, in Eq. (2) given by

$$\varphi_0=\tan^{-1}\{[\omega_0L_0-(\omega_0C)^{-1}]/R\}. \tag{12}$$

To solve Eq. (9) for i_1, we put

$$L\simeq L_0+L_1x=L_0+L_1x_0\sin\omega_1 t \tag{13}$$

in the spirit of the approximation we are using throughout and obtain

$$(L_0di_1/dt)+Ri_1+(q_1/C)=I_0L_1x_0(\omega_0\sin\omega_0 t\sin\omega_1 t -\omega_1\cos\omega_0 t\cos\omega_1 t). \tag{14}$$

Thus, we see that the contribution of i_1 amounts to an amplitude modulation of the unperturbed current with frequency, ω_1, of the oscillating pendulum.

By rewriting Eq. (14) in terms of complex functions the right-hand side becomes equal to the real part of

$$\tfrac{1}{2}I_0L_1x_0\{(\omega_0-\omega_1)\exp[j(\omega_0-\omega_1)t] -(\omega_0+\omega_1)\exp[j(\omega_0+\omega_1)t]\}. \tag{15}$$

Putting

$$\mathbf{i}_1=\mathbf{I}_1^+\exp[j(\omega_0+\omega_1)t] +\mathbf{I}_1^-\exp[j(\omega_0-\omega_1)t] \tag{16}$$

into Eq. (14), we get

$$\mathbf{I}_1^+=-I_0L_1x_0[(\omega_0+\omega_1)/Z^+]\exp(-j\varphi^+) \tag{17}$$

$$\mathbf{I}_1^-=I_0L_1x_0[(\omega_0-\omega_1)/Z^-]\exp(-j\varphi^-) \tag{18}$$

with

$$Z^{\pm}=\{R^2+[(\omega_0\pm\omega_1)L_0-((\omega_0\pm\omega_1)C)^{-1}]^2\}^{\frac{1}{2}} \quad (19)$$

and

$$\varphi^{\pm}=\tan^{-1}\{[(\omega_0\pm\omega_1)L_0 -((\omega_0\pm\omega_1)C)^{-1}]/R\}. \quad (20)$$

Thus, taking the real part again,

$$i_1=\tfrac{1}{2}I_0L_1x_0\{[(\omega_0-\omega_1)/Z^-]\cos[(\omega_0-\omega_1)t-\varphi^-] -[(\omega_0+\omega_1)/Z^+]\cos[(\omega_0+\omega_1)t-\varphi^+]\}. \quad (21)$$

Since $dx/dt\sim\cos\omega_1 t_1$ and $|\varphi^{\pm}|\leqslant\frac{1}{2}\pi$, one sees from Eq. (21) that in $F(t)$, given by Eq. (4), i_1^- will contribute a term sustaining the motion of the pendulum, whereas i_1^+ will contribute a damping term. The condition for maintenance will, then, be that the former term be larger than the latter. To obtain a quantitative relation, we note that what is relevant is the *average* of i^2 over a period of the impressed emf with frequency ω_0. Denoting this average by $\langle i^2\rangle_{\omega_0}$, we have to the order of our approximation

$$\langle i^2\rangle_{\omega_0}\simeq\langle i_0^2\rangle_{\omega_0}+2\langle i_0 i_1\rangle_{\omega_0}=\tfrac{1}{2}I_0^2\{1+L_1x_0[((\omega_0-\omega_1)/Z^-)\cos(\omega_1t+\varphi^-)-((\omega_0+\omega_1)/Z^+) \times\cos(\omega_1t-\varphi^+)]\}. \quad (22)$$

Substituting into Eq. (22) from Eq. (5) and its derivative, we finally obtain for the average magnetic force Eq. (4):

$$\langle F(t)\rangle_{\omega_0}=\tfrac{1}{4}I_0^2L_1-\tfrac{1}{4}I_0^2L_1^2\left[\frac{\omega_0+\omega_1}{Z^+}\sin\varphi^+ +\frac{\omega_0-\omega_1}{Z^-}\sin\varphi^-\right]x+\tfrac{1}{4}(I_0^2L_1^2/\omega_1) \times\left[\frac{\omega_0-\omega_1}{Z^-}\cos\varphi^--\frac{\omega_0+\omega_1}{Z^+}\cos\varphi^+\right]\frac{dx}{dt}. \quad (23)$$

The first term on the right of Eq. (23) is constant and results in a displacement of the equilibrium of the pendulum. Comparing Eqs. (23) and (3), we see that this displacement is given by

$$X=\tfrac{1}{4}I_0^2L_1/k. \quad (24)$$

The second term is proportional to the displacement of the pendulum from its equilibrium position and will thus result in a change in the frequency of the pendulum from its value when there is no current flowing.

The third term, finally, is proportional to the pendulum velocity. The condition that this term sustain rather than damp the motion is

$$0<\frac{\omega_0-\omega_1}{Z^-}\cos\varphi^--\frac{\omega_0+\omega_1}{Z^+}\cos\varphi^+ =R\left(\frac{\omega_0-\omega_1}{(Z^-)^2}-\frac{\omega_0+\omega_1}{(Z^+)^2}\right). \quad (25)$$

In view of the condition Eq. (6) this is roughly equivalent to

$$0>R\left(\frac{d}{d\omega}\frac{\omega}{Z^2}\right)_{\omega_0} =(R/Z_0^4)\left[Z_0^2-2\left(\omega_0^2L_0^2-\frac{1}{\omega_0^2C^2}\right)\right]. \quad (26)$$

We see that this condition requires a circuit with preponderantly inductive reactance and sufficiently high Q at ω_0. In particular, it will be satisfied by a circuit without capacitance if $\omega_0L>R$.

Condition (26) also implies $\varphi^{\pm}\gtrless 0$, so that the second term in Eq. (23) will be negative, and thus the frequency ω_1 will be larger than the pendulum frequency without the presence of a current in the coil.

In conclusion it should be stressed again that these results are valid only in the limit of very small amplitude of the pendulum oscillation. At larger amplitudes one cannot neglect higher order corrections to the value of the current, nor can one neglect deviations from the linear dependence of the self-inductance of the circuit on x. Then the problem becomes definitely nonlinear, and the motion of the pendulum is no longer simple harmonic (quite apart from the fact that, of course, even the unperturbed pendulum is simple harmonic only at small amplitudes).

NOTES AND DISCUSSION

Constructive and Destructive Interference

GRANT O. GALE
Grinnell College, Grinnell, Iowa

ONE of the usual demonstrations of the interference of sound waves is to send waves from a single source through a divided tube so that they unite after passing over different paths. This arrangement was first suggested by Quincke,[1] using a path difference of one-half the wavelength.

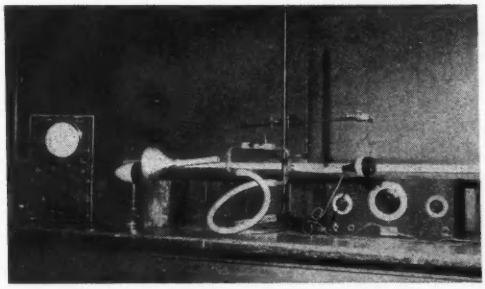


FIG. 1. Interference demonstration.

Most present textbooks[2] suggest placing the ear to the tube and when the path difference is an odd number of half-wavelengths "the sounds reunite to produce silence" (quoted from an introductory text). Even Sutton's book *Demonstration Experiments in Physics* suggests, in experiment S-103, using a whistle as the source and the ear, or a suitable pickup and audio amplifier. This does not make a satisfactory group demonstration for several reasons, notably that neither the frequency nor the length of the path difference can be readily changed. Also, the class is not as surprised as it should be when it "hears silence"!

A much more convincing demonstration is to substitute an audio oscillator for the source and a cathode-ray oscilloscope for the pickup as shown in Fig. 1. I have found the visual impression to be better than the audio because it is more quantitative.

Constructive interference is usually demonstrated first, as it is what the class expects. Since the path difference is fixed, the oscillator is tuned so that the path difference is an even number of half-wavelengths, and the volume and amplification are adjusted to give a sizable deflection on the oscilloscope. One path is then cut off by pinching the lower rubber tube, and the intensity is cut by more than half. In our setup the path difference is approximately two feet so that reinforcement occurs at about 520 vps or one wavelength. By cutting out the lower path, the amplitude on the scope drops from 1.2 in. to 0.3 in.

With the same amplification on the oscilloscope, the frequency is varied and it can be easily shown that the divided path acts as an acoustic filter when the path difference is an odd number of half-wavelengths. Thus, with the same setup, but a frequency of approximately 400 vps, only a small amplitude (0.3 in.) is observable on the oscilloscope. Now pinching the same tube increases the amplitude to about 1.5 in. It thus becomes quite obvious that whatever was going through the lower tube was actually neutralizing that which was going through the upper tube. The idea of "reinforcement" by cutting out one of the paths is not quickly forgotten by the students.

I have found the demonstration most helpful in getting across the ideas of destructive as well as constructive interference particularly in the treatment of acoustic "dead spots." It is not always clear to the students how the sound can be made more intense by cutting out one path. Also, some of our textbook authors need to be more careful in their use of technical terms. Quoting from one of them, "The phenomenon of beats observed in music illustrates destructive interference of sound waves." Query: where do the "swells" come in?

[1] See Lord Rayleigh, *Theory of Sound* (The MacMillan Company, London, 1929), Ed. 2, Vol. II, p. 63.
[2] For example, Heil and Bennett, *Fundamental Principles of Physics*, p. 322.

LETTERS TO THE EDITOR

Problems

HERE are three problems which first-year physical science students should be able to answer, but some of the graduate students might get stumped, also.

(1) A passenger riding a bus which comes to a sudden stop jerks *backward*. Explain.

(2) A wheel of radius r and weight w is rolled over a curb of height $h<r$. A student is asked to find the least horizontal force F on the axis which will cause the wheel to be on the point of rolling over the curb. His results are

$$F=w(2rh-h)^{\frac{1}{2}}/(r-h)^2.$$

By what simple process can he show that this result is incorrect?

(3) A photographer complains that he does not obtain the same exposure with his camera if he *stops down* to a certain diaphragm opening as when he *opens* to the same value. Are there grounds for this complaint?

ROBERT F. CLOTHIER
Alabama Polytechnic Institute, Auburn, Alabama

The Lumen

A CURSORY examination of elementary physics textbooks shows a great variation in the definitions of fundamental quantities. This sometimes causes confusion in the mind of a student who consults more than one book, even though the definitions are all equivalent. It is felt that certain ways of defining a quantity have greater clarity than others from the viewpoint of the student.

The following definitions of a *lumen* were taken from standard elementary textbooks. The definition from *Book A* seems to cause the least confusion in the mind of the student and is recommended as a standard definition for elementary discussions. This definition encompasses the idea of area and also rate of flow, two things that are apt to confuse the student in his thoughts about the lumen. Also the idea of unit area at unit distance makes it easy for the teacher to show the relation between the lumen and the lux.

Book A. A lumen is defined as the rate at which light energy is falling normally on a unit surface at a unit distance from a one candle power source.

Book B. A lumen is the amount of light energy falling on one square foot of a spherical surface when a point source with a strength of one candle power is placed at the center of a sphere of one foot radius.

Book C. A lumen is the amount of light flowing through a solid angle of one radian from a standard candle.

Book D. A lumen is defined as one six-hundred-and-eighty-fifth of a watt of radiant flux of wavelength 555 millimicrons, or its equivalent in evoking the sensation of brightness.

Book E. For a normal observer the lumen is equivalent to one six-hundred-and-twenty-first of a watt at the wavelength of 5550 angstrom units.

Book F. A source which is radiating with unit intensity in all directions is said to emit 4π lumens.

Book G. A lumen is defined as the light flux from a uniform point source of one candle power into a unit solid angle.

Book H. At a wavelength of 5500 angstrom units each watt of radiation corresponds to a light flux of 685 lumens.

Book I. If the radial flux in a sample of flux of wavelength 555 millimicrons is one watt, the luminous flux is 685 lumens.

Book J. If a source of light measuring one candle power is at the center of a sphere of one foot radius, the amount of light which falls on each square foot of its spherical surface is a lumen.

Book K. A lumen is the flux emitted by a point source of one candle through a solid angle of one steradian.

Book L. If we have a one candle power lamp illuminating a surface of one square foot at a distance of one foot, there will be one unit of luminous flux impinging upon the surface and this is called a lumen.

Book M. The amount of light energy emitted in one steradian by a standard candle is called a lumen.

Book N. A perfectly diffuse surface whose luminance B is one candle per square centimeter has a luminous emittance L of π lumens per square centimeter.

HARRY PEACH

Brooklyn College,
Brooklyn 10, New York

Simple Dynamics and the Value of g

LET us assume that the earth is a uniform sphere without orbital motion or rotation: Let $\gamma =$ the attraction of the earth for a unit of mass on its surface; γ is the same all over the earth. The earth, however, rotates once about its axis in 86164 sec; and to supply the necessary centripetal force on a particle at rest on the surface of the earth some small part of the γ is used up, and the remainder gives the resultant gravitation attraction which we call g. Another consequence of this is that a plumb line does not point to the center of the earth.

The relation between γ and the value of g at any latitude λ, say g_λ, is proved in many textbooks (e.g., Wagstaff's *Properties of Matter*). It is $g_\lambda = \gamma - \omega^2 R \cos^2\lambda$, where $R =$ the radius of the earth and $\omega =$ the angular velocity of rotation. In cgs units, if we assume a value of g at the poles of 982.00 cm sec^{-2} (and this is also the value of γ), then $g_\lambda = (982.00 - 3.39 \cos^2\lambda)$ cm sec^{-2}.

The center of the earth being assumed at rest (i.e., neglecting any orbital motion) the kinetic energy of a particle of mass m on the surface of the earth in latitude λ is $\frac{1}{2}m\omega^2(R\cos\lambda)^2$. If this particle is lifted a height h (h being small in comparison with R) and put on a bracket fixed to the earth, its kinetic energy becomes $\frac{1}{2}m\omega^2(R+h)^2\cos^2\lambda$. Thus, its increase of kinetic energy $= \frac{1}{2}m\omega^2(R+h)^2 - R^2\cos^2\lambda = m\omega^2 Rh\cos^2\lambda$, very nearly $= m(\gamma - g_\lambda)h$.

Similarly, if a particle of mass m originally on a bracket at height h falls into a pocket on the surface of the earth, its kinetic energy is decreased by $m(\gamma - g_\lambda)h$. Its potential energy decreases by $mg_\lambda h$. Therefore, the total loss of mechanical energy (which may be turned into heat by impact in the pocket) $= m(\gamma - g_\lambda)h + mg_\lambda h = m\gamma h$, and this is the same all over the surface of the spherical earth. If this reasoning is correct, the result is interesting and possibly new.

At the equator $g = 978.61$ cm sec^{-2}, and at every place $\gamma = 982.00$ cm sec^{-2} (assumed value, see above). Therefore, the difference between the equatorial value of g and the value of γ is 3.39 cm sec^{-2}. This difference is of the order of 0.3 percent. So that, if the heating effect produced by the impact of a falling body at the equator is calculated from the equation, energy available $= mgh$, it is in error by this amount.

In the above we have assumed a spherical earth. Modifications may be made to suit the actual ellipsoidal earth.

JOHN SATTERLY

University of Toronto,
Toronto 5, Canada

Who Writes for the Science Teachers' Magazines?*

SOME months ago the writer was invited to prepare a paper for a magazine intended for teachers in secondary schools. Both magazine and editor were unknown to me, so I was curious to learn why the editor had to solicit strangers, instead of being swamped by manuscripts submitted by his readers. The sample copy of the magazine showed a dozen articles—only one of which was written by a secondary school teacher. Half of the articles

were written by government and industrial scientists—that is, by nonteachers.

Since it appeared that I should not seem out of place, I agreed to submit an article. This has given me the opportunity to examine other issues, in which the authorship follows the same distribution as in the first issue I saw. I became curious to see to what extent this was typical of other magazines for secondary school teachers. It also seemed to me that the membership of the AAPT, which has in the past been concerned with the problem of the training of high school teachers of physics—only to learn that there is no such thing as a teacher of physics in the typical high school—might be interested to find out to what extent high school teachers write for their own professional journals.

I examined the current (since June, 1950) issues of three such magazines. Articles less than one page in length were ignored. In most cases an institutional connection is given for the author; when only a street address is given, I assume the author to be a teacher in a high or elementary school. Joint articles are assigned to the high school level if one author appears to hold a secondary school position. Authors are classified into five broad classes: *T* (elementary or high school teacher or supervisor); *E* (professor of education, professor in a teachers college, employee of federal or other department of education); *C* (college or university professor, other than in education or in a teachers college); *N* (nonteacher, usually employed by government or industry); and *M* (a few unclassifiable at first glance).

The results are shown in Table I.

It is evident that these magazines vary enormously in this respect. It should be added that they show rather little variation from issue to issue; thus magazine *C* had 8, 8, 9, and 7 articles in the *T* classification in the four issues examined. Magazine *A* is not the organ of any organization of teachers; *B* is the organ of an association which purports to include teachers of science from elementary school through the university; *C* is the organ of a high school teachers association. It is ironical that *B* shows the least participation by regular college teachers, although when categories *E* and *C* are combined it leads. One journal seems to think that the high school teacher needs articles by nonteachers; another, that articles by professors concerned with teacher training is of great importance; the third, that high school teachers can to a large extent write their own magazine. When we combine categories *E* and *C*, it appears that in all cases at least one third of the articles written specifically for high school teachers are written by college teachers. (Only a few articles under *E* are not written by teachers.)

No attempt has been made to conduct a thorough study. The analysis might have been made by author instead of article (joint papers are rather frequent); some weighting, either by length or by evaluation of subject matter might have been made. There may be a correlation between type of subject matter and author classification. One interesting but lengthy investigation could look for any differences between the early years of the journal's existence and its later years. There is a question whether other disciplines than science would yield similar results.

Possibly someone will be inspired to make an analysis of our own *American Journal of Physics*, to identify contributors who are primarily graduate level teachers, teachers of other subjects, nonteachers, etc.

My only conclusion from all this is that in science, a great deal of the writing for high school teachers magazines is done by persons who are not themselves high school teachers, nor even directly concerned with the training of those teachers. I am not sure whether that is a bad situation, but it was not what I would have supposed before I began this study.

TABLE I. Authorship of magazine articles in percent.

Magazine	*T*	*E*	*C*	*N*	*M*	No. of articles
A	3	8	27	60	3	37, in 3 issues
B	36	40	7	18	0	28, in 3 issues
C	54	14	24	5	3	59, in 4 issues

ROBERT S. SHAW

The City College,
New York, New York

* Based upon a paper delivered at the twentieth annual meeting of the AAPT, Barnard College, New York, February 3, 1951.

Synthetic Rutile

ON p. 386, *Demonstration Experiments in Physics*,[1] a demonstration of the reflections from diamond is described. "In the beam," says Sutton, "is placed as large a diamond as procurable. . . ."

Those finding it difficult to procure large diamonds may be relieved to hear that they may produce far more beautiful and interesting patterns with gem stones of synthetic rutile.[2] These stones may currently be purchased for eight or nine dollars per carat. The one used by the author—borrowed from the collection of our Geology Department—weighed 3.2 carats.

In addition to its being more readily procurable in large size, the superiority of this gem over diamond for the purpose of this demonstration is twofold. Whereas diamond is isotropic, rutile is a positive uniaxial crystal of extremely high birefringence ($\mu_o = 2.6$; $\mu_e = 2.9$). The rotation of a polaroid disk in the path of the light illuminating the gem thus permits one to distinguish the double pattern of internal reflections and the pattern of external reflections. Rutile's other superiority is its high dispersion, making for a far more colorful pattern than is obtainable with diamond. Dispersion of its ordinary ray is reported to be six times that of a diamond!

WILSON W. WOODCOCK, JR.

Hunter College,
New York 21, New York

[1] R. M. Sutton (McGraw-Hill Book Company, Inc., New York, 1938).
[2] C. H. Moore, Jr. Mining Transaction **184**, 194 (1949).

A Laboratory Manual to be Produced as a Memorial to Lloyd William Taylor

FOLLOWING the untimely death in 1948 of Lloyd William Taylor, Professor of Physics at Oberlin College, the American Association of Physics Teachers

decided to commemorate in some manner his many years of service to the Association and to the teaching of physics. The task of determining what form that memorial should take was assigned to a committee of which Duane Roller was chairman. After consulting many members of the Association this committee chose a *Laboratory Manual of Advanced Undergraduate Experiments in Physics* as the most suitable, and the present committee was appointed in February, 1950 to work out the details of this project and to set up the machinery needed to carry it out.

The first part of this task is now fairly complete. It is planned to have this book provide the teacher with comprehensive programs of advanced undergraduate experiments in each field of physics. The major content of the book will be new experiments, obtained largely as contributions by teachers of physics from their own laboratory experience. Experiments which may be found in easily available sources, such as well-known laboratory manuals, are to be included by reference only, with just enough description to identify them. Occasionally critical comments, such as desirable modifications, may be added. Since differences exist in the equipment available, as well as in the ideas of individual instructors, it is planned to present experiments in a form general enough to permit modifications to meet such differences. The detailed instructions needed by the students to fit the experiments to local conditions must be provided by their instructor. In this sense the book is aimed primarily at the teacher rather than the student. On the other hand it is expected that the student will find this book of valuable assistance to a clear understanding of his experimental work, and a helpful guide to any independent work which he may wish to undertake.

At the annual meeting of the Association in February of 1951 Joseph D. Elder was appointed Editor-in-Chief for this project. Among his qualifications might be mentioned here his experience as assistant editor of the *American Journal of Physics*, and as a member of the board of editors of *Demonstration Experiments in Physics*. At the same time the Association appropriated funds to support the work of the editor so that he could proceed as rapidly as possible. It is now up to the membership of the Association to support this work with contributions of ideas. This is a cooperative venture, and its success will depend upon the enthusiasm with which all of us support it.

As has been stated in a brief note in the March issue of the *Journal*, contributions may include new experiments, improvements upon standard experiments, recommendations concerning experiments published in past issues of this or other journals (including your own experience with them), lists of standard experiments which should be included in the manner stated above, and areas in which new experiments appear to be needed. It must be kept in mind that the experiments must be suitable for the *advanced undergraduate* laboratory, such as is sometimes conducted as a separate course, sometimes as a part of the courses in physics which make up the junior-senior program for physics majors. It is expected that many of the experiments may be performed with the equipment ordinarily found in such laboratories, or which may be constructed by students or the instructor, but contributions need not be limited to this category. Where unusual equipment is required it is advisable to be specific concerning how and where to obtain it, or how to have it made. Many standard experiments in the older fields of physics need revision, or replacement by others made possible by modern instrumentation. For all experiments it is desirable to point out the pitfalls to be avoided, the sources of error, and the accuracy to be reasonably expected. After you have contributed your own ideas, urge your friends to send in theirs which you admire. All contributions should be sent to Joseph D. Elder, Harvard University Press, 44 Francis Avenue, Cambridge 38, Massachusetts.

THE COMMITTEE
J. D. ELDER
T. H. OSGOOD
R. R. PALMER
DUANE ROLLER
C. W. SHERWIN
T. B. BROWN, *Chairman*

AMERICAN JOURNAL OF PHYSICS VOLUME 19, NUMBER 5 MAY, 1951

ANNOUNCEMENTS AND NEWS

John Wesley Hornbeck, 1881–1951

THE death on February 27, 1951, of DR. JOHN W. HORNBECK, recipient of the 1951 Oersted Medal and for twenty-five years head of the Physics Department at Kalamazoo College, is a loss keenly felt by the many students whom he has inspired and by the associates and friends who have been privileged to know him.

John Hornbeck, born in Exeter, Illinois, in 1881, spent early years as a sawyer before deciding to enter Illinois Wesleyan University, from which he was graduated in 1906. Graduate study culminated in 1913 in a doctorate at the University of Illinois. Marriage in 1915 to Frances Wolfe was the beginning of a long and happy family life. His teaching career started as principal of schools in Summer Hill, Illinois. He served as instructor of mathematics at Park College Academy from 1906 to 1908, as instructor of physics at Illinois in 1911–1912, and as a member of the Carleton College faculty from 1913 to 1924. After a year in research work at Cornell University, he moved to Michigan as chairman of the Physics Department at Kalamazoo College, to which he devoted a rich quarter century.

Those years under his leadership saw the expansion of facilities from two basement rooms to two floors of a modern science hall with the finest of equipment. Science

JOHN WESLEY HORNBECK

traditions grew too, until, according to the 1947 Steelman Report to the President, "Science and Public Policy," Kalamazoo College ranked 4th in the nation in percentage of its graduates going on to obtain doctorates in science. More than half of the physics majors under Dr. Hornbeck have gone on for graduate study, an enviable record for any teacher. It is particularly appropriate that, in what was to be the last month of his life, he received the Oersted Medal of the American Association of Physics Teachers, awarded each year for outstanding contributions to the teaching of physics.

But the contribution of Dr. Hornbeck to the lives of those who have known him cannot be indicated by these facts, for the impact of a great teacher is chiefly through his personality. In his exacting standards and his emphasis on accurate thinking and comprehension, he never lost sight of the individual student. In his last sophomore laboratory section his attention and patience were shared with just twelve students, for he believed in the close association that is possible only in small groups. Indeed, from the intimate circle of his family, his son chose physics for a career. He always gave himself completely to whatever he did. Never did his intellectual alertness diminish; in the Modern Physics book from which he taught there is written in as a footnote the results of meson-production experiments reported at meetings only a few weeks before his death. His faith in the value of liberal education was exemplified in a breadth of interest all too rare in this day when specialists often cannot converse intelligently outside their own field. The enthusiasm for astronomy that was his was shared with many students and with members of an active astronomy club; and each spring he could be seen pointing out to an interested group the arrival of a bird new to this region. He was an active and public-minded citizen of both campus and community, and for many years served as an elder in his church. Always did he feel that, far from being in conflict, science and liberal Christianity are benefited each by the other, and he found in the orderliness of the universe and the value of personality, expression of God's purpose.

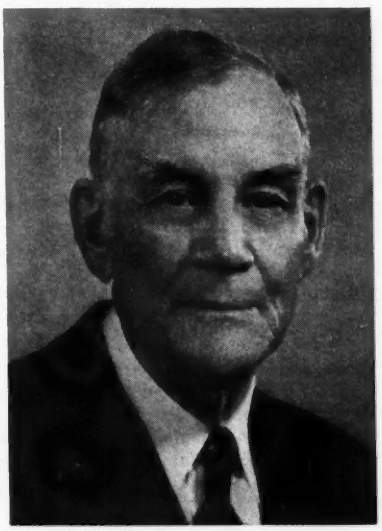
Though Dr. Hornbeck's passing is a great loss, one has a feeling that his life realized fulfillment. A scholarship fund bearing his name has been established to aid promising physics students. But the real tribute to his memory lies in the lives of those students and associates who are thankful for having known him as a gifted teacher and as a friend of great integrity, understanding, and unselfish devotion.

IAN G. BARBOUR

Louis Bevier Spinney, 1869–1951

AN association with Iowa State College begun in 1889 was terminated by the death of PROFESSOR LOUIS B. SPINNEY at Ames, Iowa, on January 25, 1951. Although ill of a heart ailment for some time, Professor Spinney continued active participation in the affairs of the Physics Department after retiring from administrative duties in 1942. The Chicago Alumni Merit Award which goes to alumni for distinguished service to the college was bestowed upon Professor Spinney in 1944. Iowa State College awarded him the honorary degree of Doctor of Science in 1948.

Professor Spinney was born at Bradford, Illinois, on May 27, 1869. In 1876 his family removed to Cass County, Iowa, where he grew up on a farm. After matriculating at Iowa State College in 1889 Professor Spinney came under the influence of W. S. Franklin, able and energetic head of the Physics Department. Professor Spinney graduated



LOUIS BEVIER SPINNEY

with the degree of Bachelor of Engineering in 1892, and in 1893 received the first degree of Bachelor of Science in Electrical Engineering to be granted by the college.

Professor Spinney began his teaching career by dividing his time between astronomy and mechanical engineering. However, inspired by Franklin he soon turned to physics teaching. After study at Cornell University and the Universities of Berlin and Zurich, Professor Spinney returned to Iowa State College, where he succeeded Franklin as head of the Department of Physics and Electrical Engineering. In 1902 he planned the quarters for electrical engineering in Engineering Hall. When physics and electrical engineering were made separate administrative units in 1907, Professor Spinney chose to continue as head of the Department of Physics. He drew plans and oversaw the building of Physics Hall, considered one of the finest structures of its kind when it was dedicated in 1922.

In his lectures Professor Spinney used no tricks of oratory or theatrical stunts. He purposely refrained from trying to do the students' thinking for them, but patiently explained and demonstrated the significance and correctness of physical principles. Professor Spinney's methods are amply illustrated in his well-known *Textbook of Physics* which appeared in five editions and sold more than 100,000 copies. At one time it was used by more than 150 colleges and universities.

Professor Spinney was ever active in alumni affairs. He served as honorary chairman of the 1950–1951 Alumni Achievement Fund. In 1945, alumni, associates, and friends presented an oil portrait of Professor Spinney to the college. For 16 years he was chairman of the college rules committee and in 1937 he served as president of the Twenty-Five Year Club. He was equally active in civic affairs.

Professor Spinney was a member of the American Physical Society, the American Institute of Electrical Engineers, the American Association for the Advancement of Science, the Illuminating Engineering Society, the Society for the Promotion of Engineering Education, and the Iowa Academy of Science, of which he was a past-president.

Professor Spinney is survived by his widow, Beryl Anna (Hoyt) Spinney, a daughter, Beryl Spinney Burns, Greenwich, Connecticut, and a son, Louis Bevier Spinney, Jr., Arvada, Colorado.

B. H. DICKINSON

Book Reviews

Technical Optics, Volume II. L. C. MARTIN. Pp. 344+vii. Pitman Publishing Corporation, New York, 1950. Price $7.50.

This volume is basically a revision of the author's earlier (1932) work *An Introduction to Applied Optics, Vol. II*, but it also contains data on some developments in optics that have occurred in the intervening years. This work uses the same style and approach as its predecessor, and hence will appeal to the same individuals who liked the earlier work.

The author develops the basic theory of geometrical and physical optics in Vol. I (reviewed elsewhere), and applies the theory to specific instruments in Vol. II. This latter volume covers representative optical devices such as telescopes, microscopes, and photographic objectives, and some related subjects.

The general approach is more descriptive than analytical, although enough mathematics to derive the fundamental theory is included. The book can be easily comprehended by a student who has assimilated the material offered in a first course in physics, and the contents of Volume I. The approach is clear, logical, and basically very sound.

The individual who has digested the contents of Vols. I and II will have a sound fundamental knowledge of optics, and will be able to handle with ease the many general optical problems encountered in academic or industrial research and development work, and will be excellently equipped to undertake the study of the design of optical systems and instruments. It is very rare that one encounters, in this country, a physicist with such a background; yet most physicists have had a reasonable number of semester hours in optics and should have progressed this far in optics. This is an unsatisfactory state of affairs, and appears to arise from a pronounced tendency to expose the physics student to a heavy dose of physical optics (a subject of some mathematical beauty, but of relatively little usefulness) rather than giving him a well-balanced dose of geometrical and physical optics, with emphasis on the interrelation of these two aspects of optics and the applicability of the subject matter to problems of all sorts. Your reviewer feels that a two-semester course in the fundamentals of optics, based on Vols. I and II of *Technical Optics* (or a similar work), would go a long way toward meeting these requirements and rounding out a budding physicist's background.

DONALD H. JACOBS
The Jacobs Instrument Company, Bethesda, Maryland

Wave Guides. H. R. L. LAMONT. Pp. 118. John Wiley and Sons, Inc., New York, 1950. Price $1.50.

This little book is an excellent review of the less technical aspects of the mathematical theory of wave guides. Although the major portion is concerned with the properties of wave guides of various shapes, the author has found room to include an adequate discussion of certain types of cavity resonators and electromagnetic horns. With but a few exceptions the problems discussed in the first six chapters are those which involve boundaries which may be identified with coordinate surfaces that separate the vector wave equation. This rather limits the scope of the book, since many of the most interesting problems in wave guide theory fall outside this category and are solvable only by various approximation methods. The present edition contains an additional chapter on wave guide techniques which partially removes the objection just stated. However, in this chapter the author contents himself with quoting results, leaving both the mathematical details and technical applications as references to be followed by the reader.

R. D. SPENCE
Michigan State College

New Members of the Association

The following persons have been made members or junior members (*J*) of the American Association of Physics Teachers since the publication of the preceding list [Am. J. Phys. 19, 255 (1951)].

Barthel, Carl C., OMR 176 Keesler A.F.B., Biloxi, Miss.
Berger, Frederick Ernest, Box 412, Lincoln Memorial University, Harrogate, Tenn.
Beyer, Norman Smith, 2152 Main St., Buffalo, N. Y.
Bowe, Jeffrey Jay, 6 Bunton St., E. Milton, Mass.
Brennan, Joseph Xavier, 7 Carmical St., Newman, Ga.
Brogan, William John (*J*), Washington College, Chestertown, Md.
Bryan, Russell Byrne, 2 N. Park St., Hanover, N. H.
Coburn, Theodore James, 224 S. Westmoor Ave., Columbus 4, Ohio.
Cole, Dandridge MacFarlan, Cilley Hall, Exeter, N. H.
Cook, Thomas Bratton, Jr., Box 1124 Vanderbilt University, Nashville, Tenn.
Coutu, Alfred Joseph (*J*), Scott Laboratory, Wesleyan University, Middletown, Conn.
Crouter, Dan E. (*J*), 822 So. Willson, Bozeman, Montana.
Crudele, Joseph Samuel (*J*), Scott Laboratory, Wesleyan University, Middletown, Conn.
Dawson, Vernon Lee, 127 Washington Ave., Hamden, Conn.
Dimmock, Thomas Herbert, 54 Gore Ct., Wake Forest, N. C.
Drake, Charles Whitney, Jr., (*J*), Scott Laboratory, Wesleyan University, Middletown, Conn.
Duncan, Ralph J., (*J*), 331 Adams Ave., Evansville, Ind.
Ellis, Jennings Bryan, New Mexico Military Institute, Roswell, N. M.
Fontheim, Ernest Gunter (*J*), 13 East Packer Ave., Bethlehem, Pa.
Gideon, Donald N., 3415 N. 24th St., Omaha, Neb.
Goldberg, Leon P., 27 Baker Ave., Dover, N. J.
Grover, Merle Fred (*J*), Apt. G-6, Stearns Village, Medford 55, Mass.
Hare, Gordon Berton (*J*), 809 Davis Ave., Takoma Park 12, Md.
Hiedemann, Egon A., Department of Physics and Astronomy, Michigan State College, East Lansing, Mich.
Hinson, Jesse J., 3105 Abell Ave., Baltimore, Md.
Hunt, Richard Allan, Department of Physics, Canisius College, Buffalo 8, N. Y.
Irvine, Merle M. (*J*), 35D East Laurel, Bethlehem, Pa.
Kane, James Joseph, Jr., (*J*), 71 Center St., Lawrence, Mass.
Kashar, Arnold Samuel (*J*), 20–69 E. 29 St., Brooklyn 29, N. Y.
King, Whitney Duncan (*J*), 805 West Pecan St., Carbondale, Ill.
Lawson, Maurice Lestel, Box 176, State College, Ark.
Lincoln, Russell Laverne, 2607 Polk Street, Corvallis, Ore.
Littleton, Melvin Edward (*J*), Washington College, Box 144, Chestertown, Md.
Longley, William Loud, St. John's School, Houston 19, Texas
Manning, Kenzel (*J*), Angwin, Calif.
McCain, Charles Edward, 15421 Stoepel St., Detroit 21, Mich.
McKinney, William Peter, Winnington Hall Club, Northwich, Cheshire, United Kingdom.
Miller, Norman B. (*J*), 153½ Vicaro Dr., Baton Rouge, La.
Missler, Herbert (*J*), 5890 Maffitt, St. Louis, Mo.
Mitler, Henri Emmanuel (*J*), 1410 Grand Concourse, New York 56, N. Y.
Muether, Herbert Robert, Department of Physics, Queens College, Flushing, N. Y.
Newman, James Blakey, 445 Institute Hill, Lexington, Va.
Okazaki, Albert (*J*), Department of Physics, University of Wisconsin, Madison, Wis.
Petty, Charles C., 250 Audrey La., Washington 20, D. C.
Quinn, William Caldwell (*J*), 1 The Circle, New Rochelle, N. Y.
Rhyins, Richard William, 143 East Darden Rd, South Bend, Ind.
Richert, Donal Eugene (*J*), Box 1562, College Place, Wash.
Rinaldo, John D. (*J*), Mile Block Rd., North Collins, N. Y.
Ruth, Ralph Powers, Department of Physics, Pennsylvania State College, State College, Pa.
Saxton, James Harrison, 5641 Monterey Rd., Los Angeles 42, Calif.
Schwenk, Harold Raymond, (*J*), 45 Lehigh St., Williston Park, N. Y.
Sellers, Ronald Earl, Jr., Box 4709, University, Ala.
Semat, Henry, 139 St. and Convent Ave., New York 31, N. Y.
Simmons, James Wood, Department of Physics, Emory University, Emory University, Ga.
Strother, G. K. (*J*), Apt. #301, 1616 Abingdon Dr., Alexandria, Va.
Thomas, James, Jr., (*J*), 4827 Willett Pkwy., Apt. 1, Chevy Chase, Md.
Townshend, Rolph, Jr., (*J*), 119 High St., Chestertown, Md.
Varwig, Robert Louis, 719 South Washington St., Fort Collins, Col.
Walker, Darcy, Laboratory of Nuclear Studies, Cornell University, Ithaca, N. Y.
Wei, Whua Fu (*J*), 424 Duncan St., Stillwater, Okla.
Werner, Ferdinand, Jr., (*J*), 4618 Chester Ave., Philadelphia 43, Pa.
Winterburg, Francis M., Veteran's Administration Hospital, Excelsior Springs, Mo.
Witte, Robert Sarnow, 2219 W. 175th St., Homewood, Ill.
Woodcock, Richard Forrest (*J*), Physics Department, Brown University, Providence, R. I.
Zechiel, Leon N., 2245 Neil Ave., Columbus 1, Ohio.

Proceedings of the American Association of Physics Teachers

The New York Meeting, February 1–2–3, 1951

THE twentieth annual meeting of the American Association of Physics Teachers was held at Barnard College and Columbia University, New York, on February 1–3, 1951. The committee for the meeting included HENRY A. BOORSE, VERNET EATON, GERTRUDE B. WERTENBAKER, and MARK W. ZEMANSKY.

Opportunity was provided for members of the Association to visit City College and Hunter College. A symposium and demonstration of recording techniques was held at the Juilliard School of Music on Saturday morning, February 3, 1951. A joint dinner with the American Physical Society was held on Friday evening in the Grand Ballroom of the Hotel New Yorker. The after-dinner address was given by J. R. OPPENHEIMER, *Institute for Advanced Study*.

The customary meeting of the Executive Committee took place on February 1, 1951, at 7:00 P.M. in the conference room of the American Institute of Physics. On February 3, 1951, in the Brinckerhoff Theater, Barnard College, the annual business meeting was held. DUANE ROLLER, *President of the Association*, presided at both meetings.

Invited Papers and Reports

Papers and Special Events

Demonstration lecture on surface tension. ERIC M. ROGERS, *Princeton University*.

Seeing light and color. LYLE BREWER substituting for RALPH M. EVANS, *Eastman Kodak Company*.—The process of seeing is somewhat different from commonly accepted notions. Three sciences are involved in the understanding of its principles and characteristics. All vision of the external world requires light. Light is a physical phenomenon and all the principles of its action are described by the science of physics. This light enters the eye of an individual and affects the nerve endings on his retina. From these an electrical current is produced which travels back to his brain. This part of the subject properly falls in the science of physiology. Certain effects produced in the brain follow well-defined laws and are quite predictable in nature. These effects are embraced by the science of psychophysics.

After a brief discussion illustrating the part played in vision by these differing types of action, the lecture was devoted to a carefully illustrated discussion of the way in which the mind interprets the information so received. A distinction is made between form and color vision and it was shown that for the most part, what we see depends as much on ourselves and our experiences as on the external reality which the light presents to our eyes. The discussion then turned to the seeing of color and, in particular, to the seeing of colored objects. By an extensive series of pictures it was shown that seeing is largely a matter of recognition of objects with properties believed to be possessed by those objects. From this it was shown that the mind has the ability to see several things simultaneously at the same spot. It follows that it is not entirely the physical or physiological facts which determine what we see but also to a great extent our knowledge of external reality as supplied by the mind. As the best example of this it was shown how it is possible for a person to see simultaneously objects illuminated by light of a certain color and at the same time see the true colors of the objects themselves. This is one of the most basic types of vision and yet it cannot be predicted from the simple physics of the light or the known properties of the eye.

Focusing sound waves with microwave lenses. WINSTON E. KOCK, *Bell Telephone Laboratories*.

The teaching of physics. W. F. G. SWANN, *Bartol Research Foundation*.

Symposium: Some problems in the recording of classical music. HARRY L. ROBIN, *Juilliard School of Music*.

Joint Session with the American Physical Society

Molecular beam researches in nuclear and electronic physics—address of the retiring President of the American Physical Society. I. I. RABI, *Columbia University*.

Oersted Medal Ceremony

Presentation of the Oersted Medal of the American Association of Physics Teachers to Professor John W. Hornbeck. DUANE ROLLER, *President of the Association*, and J. W. BUCHTA, *Chairman of the Committee on Awards*.

Some reflections on the teaching of physics. JOHN W. HORNBECK, *Kalamazoo College*.

Tenth Richtmyer Memorial Lecture of the Association

The electron theory of solids. JOHN C. SLATER, *Massachusetts Institute of Technology*.

Contributed Papers, with Abstracts

Two sessions were devoted to the following contributed papers:

1. Proposed international standards of sign conventions and symbols for geometrical optics. STANLEY S. BALLARD, *Chairman, U. S. A. National Committee of the International Commission of Optics*.—At its London meeting of July 1950, the International Commission of Optics appointed a special committee to prepare a set of proposed standards for dissemination among the national committees of the several cooperating countries, for comment by in-

terested scientists. A first draft of this set of standards is now available, and it is given below. Discussion of the various points is invited, so that the International Commission of Optics can be informed as to the attitude of American physicists toward these recommendations; comments by teachers of physics are particularly desired.

SIGN CONVENTIONS: (1) The axis of the optical system shall be the x-axis of the reference co-ordinates, except in certain cases where it is taken as the z-axis. (2) If there is no special reason to make another choice, the sense of the incident light shall be from left to right. (3) The radii of curvature of surfaces shall be measured from the poles, i.e., a surface convex towards the incident light will have a positive radius of curvature, and conversely. (4) The focal length on the object side shall be measured from the principal plane in the object space, and the focal length on the image side from the principal plane in the image space. (5) The distance from a point on the axis shall be taken as positive above the axis and as negative below the axis. SYMBOLS: (6) Gothic letters shall not be used. (7) When corresponding elements in object and image space are designated by the same letter, the elements in the image space shall be distinguished by the sign prime ($'$). (8) Reflection shall be considered as a special case of refraction, by putting $n' = -n$. (9) Points shall be designated by capital italic letters. (10) Lengths or segments shall be designated by small italic letters. (11) Angles shall be designated by small Greek letters.

2. Demonstration experiments. WALDEMAR NOLL, *Berea College.*—(*a*) Multiple real images formed by a convex lens, and (*b*) a determination of the specific gravity of a wood block using Archimedes' principle. These experiments were demonstrated at the Iowa Physics Colloquium in 1948.

3. A student spectrograph from surplus equipment. RALPH A. LORING, *University of Louisville.*—Several small spectrographs have been constructed using components usually found around a laboratory. A two-inch seasoned hardwood plank has served for a base. Slits have been constructed with fixed razor blade jaws or have been borrowed from spectrometers. The lenses have been thin meniscus lenses, rapid rectilinear camera lenses, or war surplus achromats. The dispersing elements have been plane transmission replicas, or 105-cm-radius-of-curvature, 15,000-line replicas and 60-degree glass prisms. A $3\frac{1}{4}\times 4\frac{1}{4}$-inch metal plate holder sliding in vertical ways works well at the observing end. Constructional details and details of adjustment were given.

4. Three lecture demonstrations. FRANCIS W. SEARS, *Massachusetts Institute of Technology.*—(*a*) A spring gun[1] projects a golf ball with a horizontal range of 24 ft at an angle of elevation of 45°. A paper-covered hoop through which the ball passes demonstrates that the height of the trajectory at its midpoint is one-quarter of the horizontal range. (*b*) A model of the Millikan oil-drop experiment consists of a pith-ball given lateral stability by three fine black threads. A small van de graaff generator provides an electric field which supports the ball, between horizontal plates. (*c*) The interference fringes formed by reflection are superposed on those formed by transmission, resulting in a uniform field. The film is formed between the hypotenuses of two 45°–45°–90° prisms.

[1] Fried, Green, and Mais, Am. J. Phys. 14, 1933 (1946).

5. The semidiurnal tidal oscillation of the earth's atmosphere. HAROLD L. STOLOV, *Hunter College and New York University.*—The barometric oscillation, principally semidiurnal in character, with maxima at about 10 A.M. and 10 P.M. local solar time, is interpreted according to current theory as an air-tide depending on the sun alone. Since the tide-raising force is proportional to M/d^3, where M is the mass of the disturbing body, and d its distance from the center of the earth, the moon by virtue of its proximity to the earth exerts, in spite of its smaller mass, a tidal force 2.2 times more powerful than the sun. The lunar component of the atmospheric tide should therefore be 2.2 times stronger than the solar component. This, however, is not found to be the case. Actually, what is observed is a solar semidiurnal oscillation about 100 times greater than might be expected, and a lunar semidiurnal oscillation, so feeble at middle latitudes, that it escaped detection by Laplace (1823) and Airy (1877) and was only recently demonstrated by Chapman (1918). The difficulty was resolved with a suggestion by Kelvin (1882), that the atmosphere has a period of free oscillation close to 12 solar hours. This is the famous "resonance theory," and illustrates how the effect of a small tide-generating force might be magnified. The mathematical theory of oscillations in a rotating atmosphere was briefly discussed. Linearized equations were arrived at by perturbation methods. Conclusions of "resonance theory" were presented.

6. The effect of a constant force on a particle in a box. VERNON MYERS, *Pennsylvania State College.*—Schrödinger's equation is solved for a particle in a one-dimensional "box" subject to a potential of the form kx, where k is a constant. For definiteness, an electron is assumed to oscillate vertically under the influence of gravity in a box 0.3 cm in length. The ground-state eigenfunction and energy are found, and a comparison is made between the quantum-mechanical probability distribution and the classical probability distribution for the ground state.

7. Demonstrating harmonics and beats. RICHARD C. HITCHCOCK, *Westinghouse Electric Corporation.*—Ten audio oscillators, with frequencies $110\times N$ vib/sec, where $N=1, 2, 3 \cdots 10$, can show most of the simple relations of harmonics and beats. These frequencies are chosen so that the fourth harmonic can be checked against a standard 440 pitch pipe. A phase-shift oscillator[1] gives a very good sine wave. It uses one 6SJ7 tube, 2 paper capacitors, 3 mica or "trimmer" capacitors, one electrolytic by-pass, 5 carbon resistors ($\frac{1}{2}$ watt), an output potentiometer, switch, source of heater ac and plate dc voltage. The usual audio amplifier, loudspeaker, and cathode-ray oscilloscope make up the equipment. One of the three capacitors which govern frequency is a screw-driver-adjusted trimmer, either to tune to an exact harmonic, or to give a non-harmonic tone, such as a bell. As an example of use; sound 660 and 880, the difference tone of 220 results. Then add

990, and the higher harmonic addition apparently lowers the tone by an octave. The oscilloscope images are effective in tieing together the textbook curves of sine wave addition with actual experiment.

[1] Ginzton and Hollingsworth, Proc. Inst. Radio Engrs. 29, 43–49 (Feb. 1941).

8. Selected simple demonstration experiments in mechanics. RICHARD M. SUTTON, *Haverford College.*—Four extremely simple experiments were shown to demonstrate interesting problems in mechanics in unusual ways. (1) A two-fluid siphon which interchanges two immiscible liquids of different densities and accentuates the fact that *every* siphon raises a less dense fluid to a higher elevation as it lowers a more dense fluid. (2) A simple arrangement showing the large forces introduced into a static system resembling a tight-wire rig wherein a small weight successfully lifts two 1-kilogram weights. (3) A simple dynamic arrangement to show the increase in "weight" of a pendulum bob at the bottom of its swing, and the independence of this weight on the length of the supporting string, but its dependence on the angle of swing. Explicitly, a 1-kg weight is made to raise a 2-kg weight for a 60° swing, whereas it may raise a 3 kg weight for a 90° swing. (4) Free fall of a system of two masses connected by an elastic string, and simultaneously released, as compared with a similar system connected by an inextensible string, with analysis of the accelerations and velocities involved.

9. A preprofessional orientation course for physics majors. STANLEY S. BALLARD, *Tufts College.*—The ivory tower which some of us academicians proudly claim as our part-time residence is shared also to a large extent by our students. In general, they live in a highly artificial environment, and the transition from this to the world of commerce, industry, or wherever else their chosen profession leads them, is an abrupt and perhaps unnecessarily harrowing experience for them. Therefore, at Tufts college we have been giving an orientation course which we believe helps prepare our physics majors for their entrance into the "outside world," and gives them a fuller understanding of the profession of physics. The course is conducted as a weekly seminar; the students take notes in bound notebooks which are inspected occasionally by the instructor, and they are given examinations which cover the important points that have been studied. Among the topics included are: professional and honor societies of interest to physicists, the important journals of physics and closely allied fields, the leading graduate schools in physics (of special interest to those who plan to go on to higher degrees), and a several-week discussion of "careers for physicists." This last topic is designed to give the students an insight into the various types of professional openings for physicists: academic teaching and research, government laboratory work, the broad range of positions in industry, etc. Based on our three-year experience and the letters received from recent graduates, a course of this type is not only popular with the students, but it is definitely helpful to them.

10. The development of a course "Backgrounds of Science." NORRIS W. GOLDSMITH, *Adelphi College.*—The purpose of the course is to provide, for nonscience majors, a course which will meet their needs more effectively than the standard courses offered by individual departments. The objectives are: (1) To acquaint students with the basic branches of natural science, astronomy, biology, chemistry, geology, mathematics, and physics. (2) To lead students to realize that each branch represents man's effort to observe, study, and understand one of the aspects of the universe in which he finds himself. (3) To help students to understand what has impelled men to become scientists, and how scientists approach their problems and solve them. (4) To develop an understanding of the significance of scientific thought in relation to human knowledge, the life of man, and the life of each individual. The methods used are varied. (1) Lectures are based chiefly upon experimental demonstrations and the questions and discussion which arise therefrom. (2) The topics and demonstrations for each lecture are selected by the teacher, in order to maintain a balance between the several branches of science. However, demonstrations are modified, or additional ones introduced during the lecture as need arises. (3) No attempt is made to present a large body of factual material as an end in itself, or to include in each branch of science a treatment of all of the topics which are conventionally considered essential.

11. A stop clock with magnetic fluid clutch. V. EATON, *Wesleyan University.*—A shaft, geared to a synchronous motor, makes 1 rps. This shaft is coupled to the input end of a magnetic fluid clutch,[1] and a pointer is attached to the output end. Since a friction plate prevents motion of the pointer except when the magnetizing coil is energized, the duration of the energizing current may be read to thousandths of a second from a circular scale. To check the accuracy of this apparatus as a stop clock, a falling body was used to close one micro-switch and open another. Twenty observations gave 0.3816 ± 0.0002 sec for the time of fall between the two switches. The probable error in this time measurement includes the error owing to switch operation as well as inaccuracy of the clock. When the shaft makes 10 rps, somewhat better accuracy is obtained. Improvements to increase the accuracy of the watch still further were discussed.

[1] Obtained through the generosity of the Raymond Engineering Laboratory, Inc., Middletown, Connecticut.

12. Electric discharge in air at reduced pressure. JULIUS SUMNER MILLER, *Dillard University.*—An earlier paper[1] reported an observation on a discharge-in-vacuum phenomenon. A long glass tube, arranged to show conductivity at reduced pressures, *equipped with pointed electrodes*, failed to show the usual faraday dark space. It was suggested that in a tube equipped with *pointed* electrodes, a vastly different electrical state exists than in one equipped with *plate* electrodes. This paper extends the investigation. Theory and experimental study reveal that a vastly different electrical state exists in a tube *for every different-shaped electrode.* The magnitude and gradient of the potential drop in the cathode dark space depend on the electrode shape. This gradient governs the glow characteristics, which, then, vary with the *geometry* of the electrode. Particularly, the space charge distribution de-

pends on electrode geometry, and the space charge concentration in the faraday dark space is altered by electrode shape. The usual and characteristic sequence of phenomena in such a discharge tube appears only with circular disk electrodes. Textbooks do not make this clear.

[1] Julius Sumner Miller, Am. J. Phys. 17, 448 (1949).

13. Wave fronts, attenuation, and diffraction. W. CULLEN MOORE, *Boston University*.—The concept of wave fronts as being composed of points of like phase can be explained by the use of an audio oscillator, loud-speaker, microphone, and cathode-ray oscillograph. The attenuation of intensity with distance and the shapes of wave fronts radiated from plane, line, and point sources can be qualitatively demonstrated. The double slit diffraction pattern can be explored as a function of the wavelength, slit spacing, and distance to the plane of observation.

14. The spiral program in high school physics. ALEXANDER EFRON, *Stuyvesant High School, New York City*.—This is a reorganization of the one-year college-board type of physics course, taught on a two-level cyclic basis. First-level material is semiquantitative, descriptive, and experimental; second-level material is much more rigorous and analytical. Some topics are completed on first presentation; most are spiraled and retreated mathematically. Still further spiraling takes place in honor classes. Review work is necessarily reduced, since review elements are present in all second-level instruction. Among the topics spiraled are: electrical circuits, rays and waves, units (gravitational, followed by absolute), heat calculations, statics followed by dynamics, etc. The student is made aware that additional spiraling will be required at still higher levels.

15. Who writes for the science teachers magazines? ROBERT S. SHAW, *The City College of New York*.— In the past this association has been concerned with the question of the training of teachers for the high schools. As a sort of footnote to this concern, we might consider one aspect of the professional activities of such teachers. Although the number of science teachers in secondary schools is very large, and the number of magazines published for the field is not especially large, it is surprising to note the extent to which the articles in these magazines are written either by professors in colleges or by scientists in governmental or industrial laboratories. We may speculate on the reasons for the situation, and perhaps wonder whether or not it indicates a healthy state of affairs.

Annual Report of the Treasurer

Balance brought forward from December 31, 1949			$ 5,005.01
INCOME			
Collection of dues by AIP:			
Undeposited from 1949		$ 2,874.00	
Collected in 1950:			
First Qr.	$11,424.50		
Second Qr.	1,817.00		
Third Qr.	435.00	13,676.50	$16,550.50
Royalties:			
Feb. Payment	188.62		
Aug. Payment	133.87		322.49
Bond Interest			
U. S. Treasury Bonds		300.00	
Credit from 1949			
Refund, ACE Dues		100.00	
		17,272.99	17,272.99
Total cash available, Year 1950			22,278.00
DISBURSEMENTS			
Payments to AIP			
For AJP Publication	6,609.06		
Less credit for estimate error	− 704.90	5,904.16*	
Charge for dues collection		893.92	
15% of dues collected in 1949		2,355.97	
3 month group subscription to *Physics Today*		1,138.50	
Total Paid to AIP		10,292.55	
Office Expenses and Travel			
President's office	71.28		
Secretary and Treasurer	1,091.66		
Secretary (travel)	45.00		
Editor's office	1,268.91		
Assistant Editor (salary)	800.00		
ACE Representative	39.63		
		3,316.48	
Committee Expenses			
Membership Committee	$ 1,307.75		
Nominating Committee	13.56		
Program Committee	65.96	1,387.27	
Honorarium			
Richtmyer Lecture		100.00	
ACE annual dues	100.00		
Less reimbursement from AIP	−100.00	——	
Miscellaneous	12.60		
Less postage refund (AIP)	− 2.82	9.78	
Total disbursements, Year 1950		15,106.08	15,106.08
Balance on hand, December 31, 1950			$ 7,171.92

* Applicable to fourth quarter of 1949 and first three quarters of 1950.

Securities on deposit with State Bank and Trust Co., Evanston, Illinois in a Safekeeping Account in the name of the American Association of Physics Teachers: U. S. A. Treasury Bonds, 1952–54, 2%, with par value of $15,000.00

PAUL E. KLOPSTEG, *Treasurer*

I have audited the books of account and records of Dr. Paul E. Klopsteg, Treasurer of the American Association of Physics Teachers, for the year ended December 31, 1950, and hereby certify that the foregoing statement of receipts and disbursements correctly reflects the information contained in the books of account. Receipts during the year were satisfactorily reconciled with deposits as shown on the bank statements, and all disbursements have been satisfactorily supported by vouchers or other documentary evidence. U. S. Government Bonds and Notes of $15,000 par value are held in safekeeping by the State Bank and Trust Co. of Evanston, Illinois, and a certificate was obtained from the custodian as of December 31, 1950.

WILLIAM J. LUBY
Certified Public Accountant

Chicago, Illinois
January 17, 1951

Report of the Secretary

IN summarizing the activities which made up the annual meeting, no attempt can be made to cover the program as a whole. Many of the papers presented will be published in the *Journal*. This is true also of the report of the treasurer as well as reports of committees. The following is a summary of the main business transacted at the meeting of the Executive Committee and the annual business meeting, both presided over by Dr. Duane Roller.

The following list includes all those present at the

Executive Committee meeting held February 1, from 7:00 P.M. to 12:30 A.M. at the American Institute of Physics Building: H. A. Barton, *American Institute of Physics;* C. E. Bennett, *University of Maine;* *L. Bockstahler, *Northwestern University;* *B. L. Brinker, *St. Vincent College;* T. B. Brown, *George Washington University;* *J. W. Buchta, *University of Minnesota;* B. H. Dickinson, *Michigan State College;* H. L. Dodge, *Norwich University;* R. L. Edwards, *Miami University;* J. D. Elder, *Harvard University;* R. C. Gibbs, *National Research Council;* E. Hutchisson, *Case Institute of Technology;* *H. A. Kirkpatrick, *Occidental College;* *P. Kirkpatrick, *Stanford University;* *R. A. Loring, *University of Louisville;* C. M. Loyd, *West Virginia Institute of Technology;* W. Noll, *Berea College;* *T. H. Osgood, *Michigan State College;* *R. R. Palmer, *Beloit College;* *R. F. Paton, *University of Illinois;* *W. B. Pietenpol, *University of Colorado;* *E. R. Pinkston, *U. S. Naval Academy;* *O. L. Railsback, *University of Illinois;* *D. Roller, *Wabash College;* H. K. Schilling, *Pennsylvania State College;* *F. W. Sears, *Massachusetts Institute of Technology;* *F. G. Slack, *Vanderbilt University;* M. H. Trytten, *National Research Council;* B. B. Watson, *U. S. Office of Education;* M. W. White, *Pennsylvania State College;* *J. G. Winans, *University of Wisconsin;* and *M. W. Zemansky, *City College of New York.* Names designated with an asterisk were voting members of the executive committee.

President Roller called the meeting to order at 7:00 P.M. As first order of business the report of Chairman Kirkpatrick of the committee on the constitution and by-laws was heard. Copies of the new constitution had been distributed prior to the meeting, the policy committee had previously spent some hours discussing its provisions, and the executive committee again discussed implications involved in changed procedure and policy. Finally it was moved by Professor Pietenpol of Colorado and seconded by Professor Bockstahler of Northwestern that the secretary conduct a mail ballot of the membership for adoption of this constitution and its by-laws. After further discussion this motion was carried by a vote of fourteen to two. In order to make it possible legally to conduct the membership mail ballot on both the new constitution and the new by-laws it was agreed to submit an amendment to the present by-laws at the annual business meeting. This was done and the new amendment to the by-laws was adopted. This permitted amendment of by-laws by mail ballot.

Reports of other committees were then taken up. Items of notable importance are listed. Membership in the AAPT increased by 375, representing a net increase of 275 juniors and 100 regular members. If the new constitution is adopted, junior members will have the privilege of this status for a total of five years. The treasurer's report indicated an increase in favorable balance of $2000.00 during the year. The need for a new directory of the membership was pointed out, and an appropriation to cover this expense was included in the budget. Another item of budgetary significance was the provision to cover the expense of editing of the *Taylor Memorial Laboratory Manual of Advanced Experiments in Physics.* Chairman T. B. Brown of this project reported favorable progress, requested continued cooperation of the membership in sending in suggestions, and announced that the part time services of J. D. Elder, Harvard University Press, had been secured to assist in the preparation of the text.

Reports were heard from the Editor, the AAPT representative on the American Council on Education, ACE, our member of the AAAS Cooperative Committee on Teacher Training, the representatives of AAPT on the governing board of the AIP and our representative on the committee set up to investigate the "Situation of and prospects for *Physics Today.*"

Editor Thomas H. Osgood reported on the problems and prospects of the *American Journal of Physics* and announced that his three-year appointment had terminated. The executive committee voted their appreciation of his effective services and were able to secure his acceptance of another three-year term. Dean H. K. Schilling, one of the representatives on the ACE, reported on the Washington meeting of this organization at which the problems of mobilization of scientific personnel were discussed. M. H. Trytten, H. A. Barton, and R. C. Gibbs also gave their picture of the problems confronting physicists and all scientists in the present mobilization. This report was thought so important that part of the time of the annual business meeting was assigned to Dr. Trytten, who with Dr. Barton and Dr. Gibbs reported again to a large group of the Association's membership.

Plans for the June meeting, which was agreed to be held at Michigan State College, East Lansing, Michigan, were discussed. This will be a meeting held in conjunction with the annual meeting of the American Society of Engineering Education. Plans for this meeting were already well developed, and further details will appear in the program. It was announced that C. J. Overbeck of Northwestern University had been appointed to represent AAPT on the AIP committee planning for the important anniversary meeting of AIP with all its member societies. This meeting is planned to be held in Chicago in late October this year.

At the annual business meeting the report of the committee on preparation of the new constitution was given; the budget adopted by the executive committee was announced; Professor Brown reported for the Taylor Memorial Committee; the officers of the association for 1951 were announced and the new president introduced. The new officers are

President	M. W. ZEMANSKY
Vice President	W. S. WEBB
Secretary	R. F. PATON
Treasurer	P. E. KLOPSTEG
Members of the Executive Committee	V. E. EATON
	W. C. ELMORE

The retiring president, Duane Roller, was nominated to the governing board of the American Institute of Physics to replace P. E. Klopsteg whose term expired.

A vote of appreciation was enthusiastically adopted for the courtesies extended by Barnard College to our organization and to the local committee who had planned the details so effectively.

R. F. PATON, *Secretary*